UUK. H6/5 50.53 - P0327S8

THE ACCOUNTING REVIEW

VOLUME 84, No. 5 SEPTEMBER 2009

84

JENNIFER BLOUIN, LUZI HAIL, and MICHELLE [illegible]
Capital Gains Taxes, Pricing Spreads, and Arbitrage: Evidence from Cross-Listed Firms in the U.S.

JEFFREY L. CALLEN, JOSHUA LIVNAT, and DAN SEGAL
The Impact of Earnings on the Pricing of Credit Default Swaps

JOSEPH V. CARCELLO, ANN VANSTRAELEN, and MICHAEL WILLENBORG
Rules Rather than Discretion in Audit Standards: Going-Concern Opinions in Belgium

JONG-HAG CHOI, JEONG-BON KIM, XIAOHONG LIU, and DAN A. SIMUNIC
Cross-Listing Audit Fee Premiums: Theory and Evidence

SHANA M. CLOR-PROELL
The Effects of Expected and Actual Accounting Choices on Judgments and Decisions

REBECCA FILES, EDWARD P. SWANSON, and SENYO TSE
Stealth Disclosure of Accounting Restatements

JERE R. FRANCIS and MICHAEL D. YU
Big 4 Office Size and Audit Quality

PAUL KALYTA
Accounting Discretion, Horizon Problem, and CEO Retirement Benefits

KAZUO KATO, DOUGLAS J. SKINNER, and MICHIO KUNIMURA
Management Forecasts in Japan: An Empirical Study of Forecasts that Are Effectively Mandated

W. ROBERT KNECHEL, PAUL ROUSE, and CAREN SCHELLEMAN
A Modified Audit Production Framework: Evaluating the Relative Efficiency of Audit Engagements

S. P. KOTHARI, XU LI, and JAMES E. SHORT
The Effect of Disclosures by Management, Analysts, and Business Press on Cost of Capital, Return Volatility, and Analyst Forecasts: A Study Using Content Analysis

XI (JASON) KUANG and DONALD V. MOSER
Reciprocity and the Effectiveness of Optimal Agency Contracts

A JOURNAL OF THE
AMERICAN ACCOUNTING ASSOCIATION

AMERICAN ACCOUNTING ASSOCIATION

The members of the 2009–2010 Executive Committee of the Association are as follows:

President	Nancy A. Bagranoff, Old Dominion University
President-Elect	Kevin D. Stocks, Brigham Young University
Past President	Susan F. Haka, Michigan State University
Vice President	Ira Solomon, University of Illinois at Urbana–Champaign
Vice President–Education	Belverd E. Needles, Jr., DePaul University
Vice President–Finance	Christopher J. Wolfe, Texas A&M University
Vice President–Finance Elect	Stacy E. Kovar, Kansas State University
Vice President–International	Kazuo Hiramatsu, Kwansei Gakuin University
Vice President–Professional Relations	Robert H. Colson, Grant Thornton, LLP
Vice President–Publications	Jean C. Bedard, Bentley University
Vice President–Research	Gregory B. Waymire, Emory University
Vice President–Sections and Regions	Bruce K. Behn, The University of Tennessee

The American Accounting Association was founded in 1916 as the American Association of University Instructors in Accounting. Its present name was adopted in 1936. The Association is a voluntary organization of persons interested in accounting education, research, and practice. Full membership entitles one to receive *Accounting Education News,* a selection of one, two, or all three of the Association-wide journals (*The Accounting Review, Accounting Horizons,* and *Issues in Accounting Education*) and the opportunity to purchase Association publications at a discount.

Membership in the Association is available at the following annual rates:

Full membership:
The membership dues include online selection of one, two, or all three of the following Association-wide journals: *The Accounting Review, Accounting Horizons,* and *Issues in Accounting Education.* (This does not include Section journals.)

Select one Association-wide journal	U.S. $180.00	with Print Option $230.00
Select two Association-wide journals	U.S. $190.00	with Print Option $240.00
All three Association-wide journals	U.S. $200.00	with Print Option $250.00

Student membership:
The print option is not available for student members.

Select one Association-wide journal	U.S. $25.00
Select two Association-wide journals	U.S. $35.00
All three Association-wide journals	U.S. $45.00

In addition, full members may belong to one or more Sections: Accounting, Behavior and Organizations; Accounting Programs Leadership Group; American Taxation Association; Auditing; Diversity; Financial Accounting and Reporting; Gender Issues and Worklife Balance; Government and Nonprofit; Information Systems; International Accounting; Management Accounting; Public Interest; Strategic and Emerging Technology; Teaching, Learning and Curriculum; and Two-Year College. Each Section assesses its own annual dues, which are collected by the Association.

Editorial correspondence and all manuscripts should be emailed to tar@mccombs.utexas.edu. If it is necessary to send any hard copy correspondence, the mailing address is Mary Capps, Editorial Assistant, *The Accounting Review,* The University of Texas at Austin, Department of Accounting, CBA 4M.202, 1 University Station B6400, Austin, TX 78712-0211 • Telephone: (512) 471-5318 • Fax: (512) 471-3904.

The submission fee of $200.00 for members of the AAA or $400.00 for nonmembers should be paid online (VISA or MasterCard only) at: https://aaahq.org/AAAforms/journals/tarsubmit.cfm and must be paid when article is submitted. Request for reprint permission, subscriptions, advertisements, and dues should be sent to American Accounting Association, 5717 Bessie Drive, Sarasota, FL 34233-2399 • Telephone: (941) 921-7747 • Fax: (941) 923-4093 • Email: Office@aaahq.org. Notice of change of address should be mailed at least eight weeks in advance to ensure undelayed receipt of *The Accounting Review.*

The Accounting Review is indexed in the Social Science Citation Index which is included in Web of Science.

The Accounting Review (ISSN 0001-4826 print and ISSN 1558-7967 online) is published six times a year in January, March, May, July, September, and November by the American Accounting Association, 5717 Bessie Drive, Sarasota, Florida 34233-2399. General subscriptions are available at the annual rates of $360 for a print copy or at $400 for a print and electronic copy. The member subscription rate for a hard-copy is $30. Electronic access is included with membership dues. Periodical postage paid at Sarasota, Florida and additional mailing offices. **POSTMASTER:** Send address change to *The Accounting Review,* 5717 Bessie Drive, Sarasota, Florida 34233-2399.

© 2009 American Accounting Association. All rights reserved. Printed in the United States of America.

rec'd 12\11\09

THE ACCOUNTING REVIEW

A JOURNAL OF THE

AMERICAN ACCOUNTING ASSOCIATION

Senior Editor
Steven J. Kachelmeier

Editors
Shannon W. Anderson
J. Harry Evans III
James E. Hunton
Kathryn Kadous
Sanjay G. Kallapur
Ranjani Krishnan
Laureen A. Maines
D. Paul Newman
Thomas C. Omer
Wayne B. Thomas
Mark A. Trombley
Beverly R. Walther
Paul A. Zarowin

53

VOL. 84 SEPTEMBER 2009 No. 5

JENNIFER BLOUIN, LUZI HAIL, and MICHELLE H. YETMAN
Capital Gains Taxes, Pricing Spreads, and Arbitrage: Evidence from Cross-Listed Firms in the U.S. 1321
JEFFREY L. CALLEN, JOSHUA LIVNAT, and DAN SEGAL
The Impact of Earnings on the Pricing of Credit Default Swaps 1363
JOSEPH V. CARCELLO, ANN VANSTRAELEN, and MICHAEL WILLENBORG
Rules Rather than Discretion in Audit Standards: Going-Concern Opinions in Belgium 1395
JONG-HAG CHOI, JEONG-BON KIM, XIAOHONG LIU, and DAN A. SIMUNIC
Cross-Listing Audit Fee Premiums: Theory and Evidence 1429
SHANA M. CLOR-PROELL
The Effects of Expected and Actual Accounting Choices on Judgments and Decisions 1465
REBECCA FILES, EDWARD P. SWANSON, and SENYO TSE
Stealth Disclosure of Accounting Restatements 1495
JERE R. FRANCIS and MICHAEL D. YU
Big 4 Office Size and Audit Quality 1521
PAUL KALYTA
Accounting Discretion, Horizon Problem, and CEO Retirement Benefits 1553
KAZUO KATO, DOUGLAS J. SKINNER, and MICHIO KUNIMURA
Management Forecasts in Japan: An Empirical Study of Forecasts that Are Effectively Mandated 1575
W. ROBERT KNECHEL, PAUL ROUSE, and CAREN SCHELLEMAN
A Modified Audit Production Framework: Evaluating the Relative Efficiency of Audit Engagements 1607
S. P. KOTHARI, XU LI, and JAMES E. SHORT
The Effect of Disclosures by Management, Analysts, and Business Press on Cost of Capital, Return Volatility, and Analyst Forecasts: A Study Using Content Analysis 1639
XI (JASON) KUANG and DONALD V. MOSER
Reciprocity and the Effectiveness of Optimal Agency Contracts 1671

BOOK REVIEW, Stephen A. Zeff, Editor

Martin Bloom,
Double Accounting for Goodwill: A Problem Redefined .... ALLAN BARTON 1695
Editorial Policy and Style Information 1699

Editorial Staff of *The Accounting Review*
2009–2010

SENIOR EDITOR
Steven J. Kachelmeier, The University of Texas at Austin
EDITORS
Shannon W. Anderson, Rice University
J. Harry Evans III, University of Pittsburgh
James E. Hunton, Bentley University
Kathryn Kadous, Emory University
Sanjay G. Kallapur, Indian School of Business
Ranjani Krishnan, Michigan State University
Laureen A. Maines, Indiana University
D. Paul Newman, The University of Texas at Austin
Thomas C. Omer, Texas A&M University
Wayne B. Thomas, University of Oklahoma
Mark A. Trombley, The University of Arizona
Beverly R. Walther, Northwestern University
Paul A. Zarowin, New York University
EDITORIAL ASSISTANT
Mary C. Capps, The University of Texas at Austin
RESEARCH ASSISTANT
Tracie Majors, The University of Texas at Austin

American Accounting Association

p 32,758

EXECUTIVE DIRECTOR
Tracey E. Sutherland
COMMUNICATIONS DEPARTMENT
Beverly Harrelson
Diane Hazard
Lisa Habblitz
Lindsay Wharton
Kathy Casper
Judy Cothern
Sonja Langley
Deirdre Harris
Jan Kovarik (subcontracted)
Copyediting Team (subcontracted)
OTHER AAA STAFF
Stephen Shepard
Debora Gardner
Peggy Turczyn
Barbara Brady
Pat Stein
Arlene Oliver
Kathleen Messenger
Roland LaTulip
James Szelminski
Diane Leger
Man Fong
Jodie M. Morrill
David Boynton
Alex Metz

Editorial Advisory and Review Board

Anwer S. Ahmed, Texas A&M University
Ashiq Ali, The University of Texas at Dallas
Anil Arya, The Ohio State University
Benjamin C. Ayers, University of Georgia
William R. Baber, The George Washington University
Stephen P. Baginski, University of Georgia
Ramji Balakrishnan, The University of Iowa
Tim Baldenius, Columbia University
Linda Smith Bamber, University of Georgia
Orie E. Barron, The Pennsylvania State University
Anne L. Beatty, The Ohio State University
Jean C. Bedard, Bentley University
Robert J. Bloomfield, Cornell University
Jennifer L. Blouin, University of Pennsylvania
Mark T. Bradshaw, Boston College
Lawrence D. Brown, Georgia State University
Brian J. Bushee, University of Pennsylvania
Jeffrey L. Callen, University of Toronto
Joseph V. Carcello, The University of Tennessee
Mary Ellen Carter, Boston College
Paul K. Chaney, Vanderbilt University
Qiang Cheng, University of Wisconsin–Madison
Michael B. Clement, The University of Texas at Austin
William M. Cready, The University of Texas at Dallas
Antonio Davila, IESE Business School–University of Navarra
Patricia M. Dechow, University of California, Berkeley
Mark L. DeFond, University of Southern California
Hemang Desai, Southern Methodist University
Sunil Dutta, University of California, Berkeley
Leslie G. Eldenburg, The University of Arizona
Patricia M. Fairfield, Georgetown University
David B. Farber, University of Missouri
Paul E. Fischer, The Pennsylvania State University
Jennifer Francis, Duke University
Jere R. Francis, University of Missouri–Columbia
James R. Frederickson, Melbourne Business School
Frank B. Gigler, University of Minnesota
Cristi A. Gleason, The University of Iowa
Jonathan C. Glover, Carnegie Mellon University
Steven M. Glover, Brigham Young University
David A. Guenther, University of Oregon
Luzi Hail, University of Pennsylvania
Jeffrey Hales, Georgia Institute of Technology
John R. M. Hand, The University of North Carolina at Chapel Hill
Michelle Hanlon, University of Michigan
R. Lynn Hannan, Georgia State University
D. Eric Hirst, The University of Texas at Austin
Frank D. Hodge, University of Washington
Chris E. Hogan, Michigan State University
Richard W. Houston, The University of Alabama
Raffi J. Indjejikian, University of Michigan
Christopher D. Ittner, University of Pennsylvania
Karla M. Johnstone, University of Wisconsin–Madison
Bin Ke, The Pennsylvania State University

/

S. Jane Kennedy, University of Washington
Inder K. Khurana, University of Missouri
Michael D. Kimbrough, Harvard University
Ronald R. King, Washington University in St. Louis
April Klein, New York University
Lisa L. Koonce, The University of Texas at Austin
Susan D. Krische, University of Illinois at Urbana–Champaign
Jayanthi Krishnan, Temple University
William N. Lanen, University of Michigan
David F. Larcker, Stanford University
Carolyn B. Levine, Carnegie Mellon University
Robert Libby, Cornell University
Joan L. Luft, Michigan State University
Anne M. Magro, George Mason University
Carol Marquardt, Baruch College–CUNY
Dawn A. Matsumoto, University of Washington
Edward Maydew, The University of North Carolina at Chapel Hill
Brian W. Mayhew, University of Wisconsin–Madison
Sarah E. McVay, University of Utah
Kenneth A. Merchant, University of Southern California
Gregory S. Miller, University of Michigan
Lillian F. Mills, The University of Texas at Austin
Brian Mittendorf, Yale University
Partha S. Mohanram, Columbia University
Donald V. Moser, University of Pittsburgh
James N. Myers, University of Arkansas
Venky Nagar, University of Michigan
Terry L. Neal, The University of Tennessee
Karen K. Nelson, Rice University
Edward F. O'Donnell, The University of Kansas
Per M. Olsson, Duke University
Mark E. Peecher, University of Illinois at Urbana–Champaign
Sean A. Peffer, University of Kentucky
Kathy Ruby Petroni, Michigan State University
John D. Phillips, University of Connecticut
Morton Pincus, University of California, Irvine
Joseph D. Piotroski, Stanford University
Marlene A. Plumlee, University of Utah
Suresh Radhakrishnan, The University of Texas at Dallas
Kannan Raghunandan, Florida International University
Shiva Rajgopal, University of Washington
Scott Richardson, Barclays Global Investors
John R. Robinson, The University of Texas at Austin
Darren T. Roulstone, The Ohio State University
Steven E. Salterio, Queen's University
Katherine Schipper, Duke University
Karen L. Sedatole, Michigan State University
Catherine M. Shakespeare, University of Michigan
Lakshmanan Shivakumar, London Business School
K. (Shiva) Sivaramakrishnan, University of Houston
Hollis A. Skaife, University of Wisconsin–Madison
Richard G. Sloan, University of California, Berkeley
J. Reed Smith, Indiana University at Indianapolis
Mark T. Soliman, University of Washington
Geoffrey B. Sprinkle, Indiana University
Mary Harris Stanford, Texas Christian University
Phillip C. Stocken, Dartmouth College
K. R. Subramanyam, University of Southern California
Hun-Tong Tan, Nanyang Technological University
Kristy L. Towry, Emory University
Senyo Tse, Texas A&M University
Irem Tuna, London Business School
Mohan Venkatachalam, Duke University
Joseph P. Weber, Massachusetts Institute of Technology
Michael S. Wilkins, Texas A&M University
Richard H. Willis, Vanderbilt University
Arnold M. Wright, Northeastern University
Teri Lombardi Yohn, Indiana University
Xiao-Jun Zhang, University of California, Berkeley
Mark F. Zimbelman, Brigham Young University

PAST EDITORS

1926–1929	William A. Paton
1929–1943	Eric L. Kohler
1944–1947	A. C. Littleton
1948–1949	Robert L. Dixon
1950–1959	Frank P. Smith
1959–1962	Robert K. Mautz
1962–1965	Lawrence L. Vance
1965–1967	Wendell Trumbull
1968–1970	Charles H. Griffin
1971–1972	Eldon S. Hendriksen
1973–1975	Thomas F. Keller
1976–1978	Don T. DeCoster
1978–1983	Stephen A. Zeff
1983–1987	Gary L. Sundem
1987–1990	William Kinney
1990–1994	A. Rashad Abdel-khalik
1994–1997	Robert P. Magee
1997–2000	Gerald L. Salamon
2000–2002	Linda Smith Bamber
2002–2005	Terry Shevlin
2005–2008	Dan S. Dhaliwal

THE ACCOUNTING REVIEW
Vol. 84, No. 5
2009
pp. 1321–1361

American Accounting Association
DOI: 10.2308/accr.2009.84.5.1321

Capital Gains Taxes, Pricing Spreads, and Arbitrage: Evidence from Cross-Listed Firms in the U.S.

Jennifer Blouin
Luzi Hail
University of Pennsylvania

Michelle H. Yetman
University of California, Davis

ABSTRACT: We examine how shareholder-level taxes affect the contemporaneous pricing of foreign firms' U.S. cross-listed and underlying home-country securities surrounding the 1997 reduction in U.S. capital gains tax rates. Consistent with tax capitalization, we find that the performance of cross-listed shares is negatively related to dividend yield, suggesting an abnormal price increase for shares with greater anticipated taxable capital gains. Due to barriers to cross-border arbitrage, underlying home-country securities, on average, do not react during the event, creating a temporary tax-induced pricing spread. When costs of arbitrage are low, the pricing disparity quickly dissipates and home-country shares closely mirror the pricing of their cross-listed counterparts. In further tests, we are unable to document lock-in behavior, which predicates a decrease in prices attributable to a surge in volume for shares with greater accrued taxable capital gains. Overall, our findings suggest that an exogenous shock to the U.S. tax regime reverberates in international asset prices, thereby affecting foreign firms' costs of capital.

Keywords: *tax capitalization; personal taxes; American Depositary Receipts; arbitrage; cost of capital; international asset pricing.*

Data Availability: *The data are available from public sources identified in this study.*

We thank two anonymous reviewers, Stan Baiman, Brad Barber, John Core, Gus De Franco, Dan Dhaliwal, Louis Gagnon, Paul Griffin, Bob Holthausen, Bjorn Jorgensen, Mark Lang, Christian Leuz, William Moser, Mort Pincus, Jana Raedy, Sonja Rego, Doug Shackelford, Ro Verrecchia, Bob Yetman, SungWook Yoon, Ning Zhu, and workshop participants at the 2005 University of North Carolina Tax Symposium, the 2005 European Finance Association Meeting, the 2006 American Finance Association Meeting, the 2006 Midyear Meeting of the AAA International Accounting Section, the 2006 European Accounting Association Meeting, the 2007 American Accounting Association Annual Meeting, Emory University, Massachusetts Institute of Technology, Norwegian School of Economics and Business Administration, Stockholm School of Economics, University of California, Davis, University of Chicago, University of Florida, The University of Iowa, University of Notre Dame, University of Pennsylvania, and University of Toronto for helpful comments.

Editor's note: Accepted by Dan Dhaliwal.

Submitted: August 2007
Accepted: January 2009
Published Online: September 2009

I. INTRODUCTION

We examine how personal taxes affect the pricing of stocks that are simultaneously listed and traded on multiple exchanges around the globe. Specifically, we study how an unexpected reduction in U.S. capital gains taxes around the announcement of the 1997 budget accord affects the pricing of foreign firms' American Depositary Receipts and other types of cross-listed shares in the U.S. (ADRs) relative to their underlying home-country stocks.[1] In our primary analysis, we examine the role of tax capitalization in the pricing of U.S. securities and the subsequent arbitrage-driven transfer of any pricing effects into the underlying home-country securities. The combination of the theories on tax capitalization and arbitrage allows us to make inferences on how a change in U.S. tax laws could affect foreign firms' costs of equity capital.

Tax capitalization implies that if the marginal investor in the ADR is a U.S.-taxable individual, then ADR prices, on average, should react to a change in investor-level tax rates as prices reflect investors' expectations of the securities' after-tax cash flows.[2] Absent arbitrage, the ADRs' underlying home-country securities should not react because proposed changes in U.S. investor-level taxes do not alter the after-tax cash flow expectations of foreign investors. As a result, a permanent, tax-induced pricing spread between the ADR and the home-country security is created. If, on the other hand, U.S. individual or foreign investors can costlessly move between the ADR and the home-country stock, then any change in U.S. investor-level tax rates should have no impact on the pricing spread between the two securities. As investors are assumed to be indifferent between holding the ADR or the underlying home-country stock, a fundamental price increase in one security due to the capitalization of future tax benefits is immediately followed by the same price reaction in the other security. Yet, the presence of costly cross-border arbitrage will prevent the complete price convergence. That is, the forces of arbitrage should mitigate any event-related price deviations from parity, but only to the point at which the benefits of arbitrage exceed its costs. Thus, as long as the barriers to arbitrage are sufficiently low, home-country security prices should closely mirror the pricing behavior of ADRs.

To test these assertions, we employ a sample of cross-listed firms in conjunction with an event study methodology around the May 1997 budget accord, which announced the reduction of U.S. individual capital gains tax rates. Our research design is unique as it employs a difference-in-differences approach, which allows us to investigate the interplay of U.S. shareholder-level taxes and cross-border arbitrage in the formation of asset prices with identical pre-tax cash flows, yet different after-tax cash flows (i.e., the ADR and its underlying home-country security).

We start by analyzing whether the change in personal capital gains tax rates is impounded into ADR prices. Extant evidence on the valuation consequences of shareholder-level taxes is far from conclusive and has been the subject of an ongoing debate in accounting, finance, and economics (e.g., Shackelford and Shevlin 2001; Dhaliwal et al. 2005; Guenther and Sansing 2006). Conditional on finding evidence of tax capitalization in the U.S., we then consider the transfer of such valuation effects to other countries via cross-border arbitrage. As our sample comprises firms from many different countries with various institutional backgrounds, it allows us to conduct cross-sectional analyses along the barriers to arbitrage. Next, predicated on the theory of tax capitalization, we test the persistence of any positive price impact. Evidence of a sustained price increase in ADRs' home-country securities following the event is indicative of changes in U.S. tax laws affecting foreign firms' costs of capital. Finally, our unique setting also lets us shed light on investors' lock-in behavior. As our event surrounds the reduction of U.S. capital gains tax rates, lock-in should manifest in a sudden surge in the supply of ADRs, causing downward pressure on prices. Because home-country stocks should be free of this temporary supply-side shock, they serve as an ideal benchmark to gauge any abnormal price and volume reaction of the ADR.

---

[1] Hereafter, we refer to a foreign firm's U.S. cross-listed equity as "ADR," regardless of whether it is an actual American Depositary Receipt, a direct listing, or a globally or New York registered share (see Section II for institutional details).

[2] We use the term "marginal" investor, although identifying a marginal investor is not critical to our analysis. If prices are set by the "aggregate average response" of investors, then our maintained hypothesis becomes that U.S.-taxable individuals form a large enough portion of the investor pool to have an effect on price.

We find that during the five trading days surrounding the *announcement* of the tax rate cut, the performance of ADRs is significantly and negatively related to dividend yield, consistent with tax capitalization. Based on the inter-quartile range, no- or low-dividend-yield firms outperform high-dividend-yield firms by 120 basis points. On average, no significant pattern is apparent for the underlying securities in the firms' home countries, creating a tax-induced pricing spread of 87 basis points. These results are robust to various sensitivity tests.

When we assess the mitigating role of (costly) cross-border arbitrage, we find that home-country shares of firms with lower barriers to arbitrage (as measured by proxies of high trading activity and a sophisticated investor base) closely mirror the price reaction of the ADRs during the event, rendering the pricing spreads between the two securities substantially smaller or insignificant. We then show that ADR firms' home-country shares continue to significantly outperform their local, non-cross-listed peers for at least 45 days beyond the event. Moreover, the price reaction of ADRs and underlying home-country shares is associated with firms' proclivity to raise capital. Taken together, our findings suggest that the U.S. tax rate cut reverberates in the pricing of cross-listed firms' foreign securities, thereby lowering the firms' costs of capital.

The analysis of the five trading days surrounding the *effective date* of the tax rate cut reveals that prices for ADRs with positive accumulated share appreciation decrease during the lock-in week. However, home-country prices decrease as well, resulting in an insignificant pricing spread. Moreover, we find that the price reaction is not attributable to a temporary supply-side liquidity shock, as trading volume during the lock-in week, if anything, is *below* normal for appreciated stocks. Taken as a whole, our evidence does not support lock-in behavior.

This research contributes to the literature in several ways. First, we provide empirical evidence on whether and how U.S. shareholder-level taxes affect international asset prices. We find that shareholder-level taxes are capitalized into ADR prices, and, when arbitrage costs are low, these valuation effects are subsequently transferred into home-country stock prices, implying that a firm's global investor base can have an impact on its cost of equity capital. This is a new and unique finding, adding to both the literature on international cross-listings (e.g., Karolyi 1998, 2006) and the literature on the impact of taxes on cross-border transactions (see McDonald 2001; Christoffersen et al. 2005).[3]

Second, we provide direct evidence on how the interplay of personal taxes and cross-border arbitrage contributes to the existence of a tax-induced pricing spread between firms' shares traded in the U.S. and their underlying securities. Although there exists a well-established literature examining the level of pricing spreads, prior studies merely acknowledge the potential role of taxes in creating such spreads but do not empirically test for its existence (e.g., Maldonado and Saunders 1983; Rosenthal and Young 1990; Miller and Morey 1996; Froot and Dabora 1999; Gagnon and Karolyi 2004). We overcome this limitation by building on extant theories of tax capitalization and lock-in and applying them to the 1997 capital gains tax rate cut, which allow us to empirically document both the creation and the elimination of tax-induced pricing spreads under the conditions of costly arbitrage.

---

[3] In a related study, Callaghan and Barry (2003) provide evidence on the effect of dividend taxation on ADRs. We go beyond their approach and specifically investigate the implications of U.S. taxes on the pricing of ADRs *and* the underlying home-country securities.

Our third contribution is to the literature studying the relation between taxes and investor-level capital gains taxes. At a descriptive level, we extend prior findings confined to local firms within a country (e.g., Amoako-Adu et al. 1992; Lang and Shackelford 2000), and special occasions in the life of an entity, such as initial public offerings (e.g., Reese 1998; Guenther and Willenborg 1999; Ayers et al. 2003; Ellis et al. 2006) or mergers and acquisitions (e.g., Landsman and Shackelford 1995; Erickson 1998; Ayers et al. 2004, 2007) to a sample of U.S. cross-listed firms. Our paired-securities approach enables us to effectively alleviate a major concern that plagues much of the prior research, i.e., to eliminate firm-level risk factors as a potential explanation of the findings.[4]

Finally, we add to recent studies that attempt to disentangle the long-term equilibrium pricing effects (consistent with tax capitalization) from the short-term, disposition-induced pricing effects (consistent with lock-in). Lock-in behavior is predicated on tax-sensitive investors reaching their reservation price and selling en masse, thereby increasing volume and depressing prices. As our setting allows us to simultaneously examine price and volume reactions for the same security in two separate markets, our findings ultimately challenge the interpretation of concurrent studies using U.S. firms only (e.g., Cook 2007; Ayers et al. 2008; Dai et al. 2008).

In Section II, we provide background information on ADRs and the 1997 U.S. budget accord, and develop our hypotheses. In Section III, we discuss sample selection, present descriptive statistics, and outline the empirical design. Section IV reports results for the tax capitalization tests. In Section V, we present cross-sectional analyses on the forces of arbitrage and examine the persistence of tax capitalization in the U.S. and abroad. Section VI contains the analysis of lock-in behavior. Section VII concludes.

II. INSTITUTIONAL BACKGROUND AND HYPOTHESES DEVELOPMENT

Cross-Listed Equities

Foreign-incorporated firms cross-list on U.S. exchanges most commonly in the form of American Depositary Receipts or direct listings. The former are U.S. negotiable certificates representing underlying shares in a company incorporated outside the U.S. Depositary banks (e.g., Bank of New York) immobilize shares of the home-country stock with a custodian in the home country and issue U.S.-dollar-denominated depositary receipts as claims against the immobilized shares. They also convert the proportional share of dividends and other payments into U.S. dollars before transferring them to the shareholders. These securities make it easier and more cost-effective for U.S. investors to purchase shares in non-U.S. firms. Canadian firms exclusively utilize direct listings. The shares traded in the U.S. represent the identical shares traded on the Canadian exchange, and like depositary receipts, the cross-border transactions involve a transfer agent. Under the Multi-Jurisdictional Disclosure System, Canadian firms have modified SEC reporting requirements that lower the costs of direct listing (see Frost and Kinney 1996). Prior research has documented that U.S. investors own a substantial portion of ADRs and direct listings, suggesting that the marginal investor might be a U.S.-taxable individual (Callaghan and Barry 2003; Bradshaw et al. 2004; Ammer et al. 2006).

---

[4] A notable exception is Hubbard and Michaely (1997) who study the impact of the 1986 Tax Act on Citizens Utilities, which issued two securities identical in every way except that one pays a dividend and the other does not.

In the U.S., depositary receipts, direct listings, and directly owned foreign securities are taxed similarly. Our focus is on a particular tax event that reduced the U.S. individual capital gains tax rates from 28 to 20 percent, but left the dividend tax rate and other rules guiding taxation of inbound and outbound foreign investment income unchanged. Hence, tax-specific barriers to arbitrage did not change around the event, which is a nice feature of our setting.[5] With respect to capital gains, ADRs and directly owned foreign securities held by U.S.-taxable individuals are taxed at the U.S. individual capital gains tax rates, identically to U.S. equity investments (see Callaghan and Barry [2003] for an in-depth discussion of ADR taxation).

From a risk perspective, an ADR represents the same cash flows as its underlying home-country stock. Thus, with the exception of currency risk, there is no difference in the operational riskiness between the two securities.[6] To the extent that both security types are fully fungible and that markets are efficient, the prices of the two securities should equate. Pricing spreads between the two sets of securities allow for two types of arbitrage. First, an investor can engage in "convergence arbitrage" and profit from a pricing spread by taking a long (short) position in the ADR when the ADR is relatively underpriced (overpriced) as compared to the underlying security and reverse the transaction after prices have converged. Second, an investor can engage in "instrument arbitrage" and profit from a pricing spread by exchanging shares across markets (Kim et al. 2000). For example, assuming the ADR is priced lower than the home-country security, an investor would take a long position in the ADR *and,* at the same time, a short position in the underlying stock, go through the ADR conversion process (which takes about a day and is subject to a conversion fee), and then use the de-immobilized home-country shares to settle the short position.[7] Although both of these strategies are not without risk and involve (round-trip) transaction costs, we use the term "arbitrage" (similar to Pontiff 1996) to refer to taking a position in a mispriced asset and/or taking the opposite position in its related security (see also our stylized discussion of the effects of arbitrage in the Appendix).[8]

---

[5] Contrary to capital gains taxes, withholding taxes on dividends could create tax-specific barriers to arbitrage as they differ depending upon the location of the investor (i.e., U.S. or foreign) *and* the location of the security (i.e., U.S. ADR or home-country security). For example, a U.S. shareholder investing in the ADR (home-country stock) will pay the U.S. dividend tax rate (U.S. dividend tax rate with a credit for any applicable foreign withholding taxes). Differences in taxation across the two securities (i.e., the foreign withholding taxes) effectively increase the costs of arbitrage thereby impeding price convergence.

[6] In terms of corporate governance, exchange-listed ADRs (i.e., those included in our sample) are generally subject to U.S. securities laws. However, there are some cases where ADR investors are more restricted than home-country share investors in their ability to exercise their voting rights (see Cowett 2004). While this might affect the general magnitude of the pricing spread between an ADR and the underlying share, it should not affect any change in spread during our event periods.

[7] Practically speaking, this type of arbitrage is not always successful. For example, in trying to exploit pricing discrepancies, Headwaters Capital of Sausalito, California, sought to simultaneously take a long position in the ADRs of Terra Networks, a Spanish Internet portal company, in the U.S. and a short position in ordinary shares in Spain. However, at the time, Terra's ADR securities were not very liquid (i.e., they were hard to come by). The broker successfully placed the sell order in Spain for the home-country stock but was unsuccessful in purchasing the corresponding ADRs, leaving Headwaters' arbitrage incomplete (*see Wall Street Journal* 2004).

[8] We thank Jason Paltrowitz, Vice President of Broker & Institutional Marketing at the Bank of New York ADR Division, for providing institutional details regarding arbitrage between ADRs and their home-country stocks.

The 1997 Budget Accord

The event brought about a reduction from 28 to 20 percent in the long-term capital gains tax rate on U.S. investors, but no change in the dividend tax rate. With respect to tax capitalization, we must identify the date when market expectations about future capital gains tax rates changed. It is likely that before the 1998 budget discussions the market did not predict a capital gains tax reduction. Neither President Clinton's 1996 reelection campaign nor his March 1997 introduction of the 1998 budget endorsed such a reduction. In fact, in March 1997, William Archer, the House Ways and Means Committee Chair at the time, indicated no greater than a 50 percent chance that a 1997 tax bill would include such a reduction (*Tax Notes*, March 24, 1997). There was little release of information regarding the budget discussions until April 30, when information suggesting a balanced budget was imminent. Finally, on May 2, an agreement between President Clinton and Congress members to both balance the budget and reduce the capital gains tax rate was announced. Although the actual rates were not specified until August, prior research has focused on May 2 as an event date because the announcement of the budget accord provided the market with new and fairly certain information regarding a decrease in the capital gains tax rate (Lang and Shackelford 2000). In our tests, we use the five trading days from April 30 to May 6 as the tax capitalization event week.

To investigate lock-in we must identify the effective date for the reduced rates. On May 7, 1997, House Ways and Means Committee Chairman William Archer and Senate Finance Committee Chairman William Roth issued a statement announcing that "any capital gains tax cut included in the budget reconciliation bill will be with respect to deals and exchanges occurring on or after today, May 7, 1997." Once again, although the rate is not yet known, shareholders are assured that gains realized on or after May 7, 1997 will be eligible for the reduced capital gains tax rates. Thus, the period of May 7 to May 13 is the lock-in event week.

Hypotheses Development

We develop our hypothesis from two streams of research: (1) the theories of tax capitalization and lock-in (e.g., Feldstein et al. 1980; Klein 1998, 1999; Lang and Shackelford 2000), and (2) the law of one price in a world absent frictions (e.g., Lamont and Thaler 2003; Gagnon and Karolyi 2004).

First, we illustrate the intuition of how U.S. investor-level taxes (and hence changes in these taxes) could affect ADR prices using the Klein (1998, 1999) general equilibrium-pricing model. This model is sufficiently rich to test the theories of tax capitalization, lock-in, and tax irrelevancy.[9] In the Klein model, price is determined by the present value of after-tax cash flows resulting from both dividends and capital appreciation/depreciation. The impact of capital gains taxes is twofold: (1) the taxes on expected *future* capital appreciation are "capitalized" into asset prices (i.e., price is decreasing in expected tax liabilities), and (2) tax liabilities on unrealized *past* capital appreciation increase investors' pre-tax rates of return on the firm's stock (i.e., shareholders are "locked-in" causing price to be increasing in tax liabilities from unrealized capital gains).

---

[9] There exists other theory on the relation between tax rate changes and price. Using a stylized trading model, Shackelford and Verrecchia (2002) study intertemporal tax discontinuities, which result from circumstances in which different tax rates are applied depending upon an asset's holding period (e.g., long-term versus short-term capital gains). Their predictions stem from short-term liquidity shocks around earnings announcements, namely sellers' strikes by individuals with short-term capital gains who are just on the cusp of long-term qualification. However, in our setting, any potential portfolio rebalancing is attributable to only individual investors. As such, institutional investors should be available to eliminate any potential slack in supply.

Empirically, we use dividend yield as our proxy for expected *future* capital appreciation. To the extent that prices are formed based on a weighted average of after-tax cash flows stemming from capital gains and dividends (Poterba and Summers 1985), the higher the proportion paid in dividends, the lower the proportion paid in capital gains. To capture *past* capital appreciation, we employ buy-and-hold stock returns over the year leading up to the event (e.g., Klein 1999).

Second, we develop a stylized model considering the forces of arbitrage in our paired-securities setting (see the Appendix). In one extreme, the U.S. ADR market and the home-country stock market are completely separable, and no arbitrage occurs. Thus, any one-sided price change does not affect the pricing of the corresponding security, and cross-border tax differentials result in a permanent pricing spread. In the other extreme, U.S. and foreign investors can costlessly move between ADRs and home-country securities. Consequently, personal taxes do not give rise to pricing spreads across markets. Any price change in one security immediately and completely reverts or leads to a parallel move in the other security. Between these two extreme cases, there exists a continuum where costly arbitrage effectively prevents or delays prices from reaching a new equilibrium. As a result, cross-border tax differentials can induce pricing spreads between the ADR and the underlying stock, but only to the point where the costs of arbitrage exceed its benefits. Such barriers to arbitrage include transactions costs, liquidity constraints, investor sophistication, short-selling restrictions, and market inefficiencies.

Empirically, Gagnon and Karolyi (2004) show that pricing spreads rarely persist for longer than five days, suggesting that cross-border arbitrage drives prices to parity within a time period equal to our event window. Koumkwa and Susmel (2005) document similar results, although for a smaller, more focused sample of Mexican ADRs. We explain our empirical proxies for costs of arbitrage in Section V.

Tax Capitalization

Our first set of hypotheses, in alternative form, is based on the theory of tax capitalization.[10] Consistent with the marginal investor in the ADR being a U.S. individual who anticipates incurring capital gains, we expect that the announcement of a reduction in the capital gains tax rate will increase security prices because the present value of any future capital gains tax liabilities has been reduced. Yet share prices for firms expecting to pay proportionately more of their profits in capital appreciation (i.e., firms having low dividend yields) should react more positively. The underlying stock in the home country, though, is unlikely to have a U.S. individual as the marginal investor and, absent arbitrage, should exhibit no reaction.

> **H1a:** During the capitalization week, there is a significant *positive* return for the ADR, which is *decreasing* in dividend yield, but *no* significant return reaction for the underlying home-country stock.

Conditional on the expected tax rate cut being impounded in ADR prices, we next consider the direction of arbitrage. Tax capitalization implies a permanent increase in a firm's equity value and, consequently, a decrease in the firm's cost of capital (e.g., Dhaliwal et al. 2007). Thus, ADR prices should reach a new equilibrium that persists into the future.

---

[10] Our null hypothesis of tax irrelevance (Miller and Scholes 1978) follows from the presence of tax-exempt marginal investors, taxes having no impact on equity prices, or ADRs temporarily reacting, but then reverting back toward the original equilibrium price via costless arbitrage.

It follows that, if costs of arbitrage are sufficiently low, then the underlying home-country stock should quickly converge to the higher equilibrium price, leading to our second hypothesis.

> **H1b:** When costs of arbitrage are *low*, the return for the underlying stock during the capitalization week reacts similarly in size and direction to the return of the ADR.

For easily arbitraged stocks, we expect the home-country price reaction to the U.S. capital gains tax rate cut to mimic that of its ADR counterpart, resulting in a reduced or insignificant pricing spread.[11] This is consistent with U.S. tax rates affecting a foreign firm's cost of equity capital. However, when the barriers to arbitrage are high, a pricing spread should develop during the event period as these barriers impede the convergence of the home-country stock price to the new equilibrium. Although the speed of adjustment to the new pricing equilibrium varies by arbitrage costs, the persistence of any reaction does not. Therefore, we examine whether the reaction of the foreign security to the change in U.S. taxes appears to be permanent.

Lock-In Effect

Our second set of hypotheses considers lock-in behavior. If the marginal investor of the ADR is a U.S.-taxable individual, then we expect security prices to decrease on the *effective date* of the new capital gains tax rate. This prediction results from the reduced capital gains tax rate lowering ADR shareholders' reservation prices, thereby leading to a short-term, supply-side liquidity shock, causing downward pressure on prices. It follows that ADR prices for firms with proportionately higher accrued capital appreciation should react more negatively to a decrease in capital gains tax rates. Again, the underlying home-country stock is unlikely to have a U.S. individual as the marginal investor and should exhibit no reaction.

> **H2a:** During the lock-in week, there is a significant *negative* return for the ADR, which is *decreasing* in past capital appreciation, but *no* significant return reaction for the underlying home-country stock.

The lock-in effect's impact on returns stems from a surge in trading volume. Thus, ADRs with proportionately higher unrealized capital appreciation should experience the largest increase in volume because their shareholders face the greatest decreases in accrued capital gains tax liabilities, leading to our second hypothesis.[12]

> **H2b:** During the lock-in week, there is a significant *increase* in trading volume for the ADR, which is *increasing* in past capital appreciation, but *no* significant trading volume reaction for the underlying home-country stock.

Finally, we address the interplay between lock-in and cross-border arbitrage. Under lock-in, the reduction in price is nothing more than a temporary, sell-side-driven liquidity shock that should not affect the firm's cash flows. Thus, the one-sided price deviation occurs irrespective of the security's fundamentals, and the ADR price should quickly revert to the original equilibrium. No permanent valuation effect should occur. Although barriers to arbitrage can slow down such an ADR price adjustment, under lock-in, the home-country stock is not expected to move.[13]

---

[11] Hypothesis H1b is also consistent with the marginal investor of the underlying stock being a U.S.-taxable individual, which we deem unlikely. Moreover, if indeed the marginal investor abroad is subject to U.S. tax legislation, then barriers to arbitrage should not matter for the underlying stock's reaction to the tax rate cut.

[12] In comparison, tax capitalization does not take a stance on trading volume, which is why we did not formulate a volume-related hypothesis for the capitalization week.

III. DATA, DESCRIPTIVE STATISTICS, AND RESEARCH DESIGN

Sample Selection and Description

The sample selection is designed to identify ADRs and their underlying securities with sufficient trading activity in the two years surrounding the event. We draw the initial sample from the intersection of foreign firms with U.S. prices on CRSP and home-country stock prices on Datastream. First, to identify ADRs, we use CRSP share codes (first digit of SHRCD = 3). Next, we identify Canadian direct listings using the Compustat country of incorporation (FINC = 9), and match them with CRSP. Finally, we merge our set of foreign firms with available U.S. prices to the Datastream universe and confirm the inclusion of these firms in the Citibank Universal Issuance Guide (http://www.citibank.com/adr) or the Bank of New York's Complete DR Directory (http://www.adrbny.com). This yields 179,557 pairs (ADR and underlying stock) of daily returns between May 1, 1996 and April 30, 1998, representing 431 unique firms.

Next, we require that firms actively trade in the U.S. during the event (i.e., from April 30 to May 13, 1997), which reduces the sample to 318 firms with 152,955 observations.[14] We further eliminate firms that initiated their ADR listing in the U.S. after May 1, 1996, and observations with only one-sided trading (i.e., for each trading day, volume and price data must exist in the U.S. *and* at home). The first criterion is necessary because an exchange listing in the U.S. is a special event in the lifespan of a foreign corporation (e.g., Leuz 2003; Karolyi 2004). Firms cross-listing in the U.S. experience abnormal returns at the listing announcement as well as during the pre- and post-listing periods (e.g., Karolyi 1998; Errunza and Miller 2000; Sarkissian and Schill 2008; Hail and Leuz 2009). The second criterion ensures the existence of market prices for the spread calculation. To minimize sample attrition due to variation in trading days across U.S. and foreign markets (i.e., difference in market holidays), we add returns and volume from one-sided trading days to the next observation with double-sided trading data available. This adjustment is consistent with returns after non-trading days incorporating all the information that is embedded in the corresponding returns with continuous trading.[15] Finally, we require that firms have a minimum of 100 trading days (i.e., positive trading volume) during the two-year sample period. Our final sample consists of 120,094 return pairs representing 266 firms.

---

[13] Once again, the usual cautions apply. If the marginal investor in the underlying stock is a U.S.-taxable individual, which we deem unlikely, then we expect to find a similar return and volume pattern as in the ADR. Alternatively, the underlying stock price (but not trading volume) could be influenced by arbitrage activity due to differing investors' beliefs. If arbitrageurs notice a spread growing between the ADR and the home-country prices *and* believe that it was created by information rather than a short-term liquidity shock (i.e., the lock-in effect), then they could move to eliminate the spread.

[14] We define a firm as actively traded if U.S. price and volume data exist on at least four out of the ten event days. If we limit our sample to firms with consecutive trading during the event, we lose another 47 firms. Using this smaller sample does not affect our main findings. If, on the other hand, we do not impose any trading restrictions during the event period, then the sample size increases by 28 firms, and all the results are very similar to those reported.

[15] One-sided trading is of a particular concern during the capitalization week, which happens to coincide with one or more national holidays in many countries (e.g., Cinco de Mayo in Mexico, Constitution Memorial Day in Japan, or May Day Bank Holiday in the U.K.). Without the correction, we would lose many event-day observations with valid ADR returns because of missing home-country returns. Note, though, that the results throughout our study are not sensitive to this adjustment procedure.

Table 1, Panel A, provides an overview of the sample composition. For each sample country, we report the number of firms and daily observations for all sample firms and for the subsamples that pay dividends and have positive stock price appreciation. Aside from the Canadian firms (38 percent), only firms in Australia, Japan, Mexico, and the United Kingdom compose 5 percent or more of the total observations. Panel B reports descriptive statistics for our main variables of interest. *DIVYLD* and *YLD_DUM* are the continuous dividend yield (calculated as the sum of the last 12 months of dividends scaled by daily stock prices) and the dichotomous dividend-paying indicator (equal to 1 for dividend-paying firms), respectively.[16] About two-thirds of our sample firms pay dividends over the fiscal year preceding the event. *APPR* is the yearly buy-and-hold stock return computed on a daily basis. If the stock has appreciated in the year prior to the event, then *APPR_DUM* is set equal to 1, and 0 otherwise. Again, about two-thirds of our sample firms have appreciated in value.

ADR and Home-Country Returns and Abnormal Trading Volume

In our returns analyses, we use the ADR's return in the U.S. (*RET_ADR*), the return of its underlying security in the home country (*RET_HOME*) or the difference between the two (*SPREAD*) as our dependent variable. Returns are daily changes in closing stock prices denominated in local currency (dividend and split-adjusted).[17] *PDIFF* is the natural log of the ratio of the ADR price to the underlying stock price, both denominated in U.S.$.[18] *PDIFF* is intended to account for cross-autocorrelation between *RET_ADR* and *RET_HOME* and should help control for prices' tendency to revert toward parity (Gagnon and Karolyi 2004).

In the volume tests, we use a security's abnormal trading volume in the U.S. (*AVOL_ADR*), the home country (*AVOL_HOME*) or the difference between the two (*AVOL_SPREAD*) as dependent variables. We measure abnormal trading volume as daily trading volume scaled by the mean daily trading volume in the year prior to the event.[19] To reduce skewness, we transform the raw data (denominated in local currencies) using natural logarithms.

Table 1, Panel B, reports that *RET_ADR* and *RET_HOME* have respective means (medians) of 6 and 8 (0) basis points and standard deviations of about 3.5 percent. This translates into a tight *SPREAD* distribution, with 80 percent of the observations having a return spread within ±2 percent. However, we do observe some observations with extreme values. *PDIFF* has a similar distribution. While the mean (median) price differential is −116 (19) basis points, some companies are traded in the two markets at prices far from parity.[20] Since

---

[16] We scale the dividends by the daily stock price five days before the event, as we want to avoid any spurious correlations by using event-day stock prices.

[17] In contrast to Gagnon and Karolyi (2004), we do not correct our return metrics for non-synchronous trading between the U.S. and foreign markets. First, our two event windows cover a five-day period each, which allows for some spillover effects. Second, when we include lead and lag values of the independent variables in our analyses, the results (not reported) remain virtually unchanged. Finally, the inclusion of country-fixed effects should account for any systematic biases due to partially or non-overlapping trading hours.

[18] We multiply the underlying home-country stock by the bundling ratio since ADRs commonly do not map one-to-one into home-country shares. Canadian shares require no adjustment since they trade on a one-to-one basis in the U.S. For all other firms we obtain ADR bundling ratios from either the Bank of New York ADR database or the JP Morgan Global Issuance file.

[19] The results of the abnormal volume analyses are not affected if, alternatively, we use the entire two-year sample period (excluding the event) as the denominator.

[20] When we calculate average values of *PDIFF* by firm, we obtain results that closely resemble those presented by Gagnon and Karolyi (2004): our inter-quartile range is 83 basis points with extreme observations trading at a discount (premium) of up to 76 percent (77 percent) in the U.S. compared to their home-country shares.

TABLE 1
Descriptive Statistics on ADR and Home-Country Stocks

Panel A: Sample Composition by Country[a]

Country	# of Dividend-Paying Firms	# of Appreciated Firms	Total # of Firms	# of Dividend-Paying Observations	# of Appreciated Observations	Total # of Observations	Total # of Observations in % of Sample
Argentina	6	6	7	2,881	2,881	3,073	2.56
Australia	10	10	14	4,792	4,624	6,601	5.50
Brazil	0	1	1	0	475	475	0.40
Canada	34	57	104	15,516	25,037	45,841	38.17
Chile	12	11	12	5,105	4,639	5,105	4.25
China	1	0	2	479	0	958	0.80
Colombia	1	1	1	445	445	445	0.37
Denmark	2	2	2	975	975	975	0.81
Finland	1	1	1	486	486	486	0.40
France	5	5	6	2,445	2,441	2,930	2.44
Hong Kong	1	0	1	429	0	429	0.36
Indonesia	3	0	3	1,197	0	1,197	1.00
Israel	1	1	2	465	424	889	0.74
Italy	5	4	5	2,321	1,834	2,321	1.93
Japan	15	7	16	6,926	3,305	7,355	6.12
Korea	2	0	2	946	0	946	0.79
Mexico	9	12	17	4,009	5,711	7,777	6.48
The Netherlands	9	9	9	4,141	4,141	4,141	3.45
New Zealand	3	3	3	1,473	1,473	1,473	1.23
Norway	2	2	3	862	976	1,349	1.12
Peru	1	1	1	314	314	314	0.26
Philippines	1	1	1	487	487	487	0.41
Portugal	2	2	2	897	897	897	0.75
South Africa	3	2	3	1,440	953	1,440	1.20

(continued on next page)

TABLE 1 (continued)

Country	# of Dividend-Paying Firms	# of Appreciated Firms	Total # of Firms	# of Dividend-Paying Observations	# of Appreciated Observations	Total # of Observations	Total # of Observations in % of Sample
Spain	5	5	5	2,444	2,444	2,444	2.04
Sweden	5	4	6	2,235	1,939	2,693	2.24
Switzerland	1	2	2	491	790	790	0.66
United Kingdom	32	27	35	15,121	12,713	16,263	13.54
Total	172	176	266	79,322	80,404	120,094	100.00

Panel B: Summary Statistics of Returns, Abnormal Trading Volume, Market, and Firm Characteristics[b]

	Mean	Standard Deviation	1%	25%	Median	75%	99%
ADR and Home-Country Returns and Price Differences							
RET_ADR	0.06%	3.32%	−8.93%	−1.20%	0.00%	1.20%	10.00%
RET_HOME	0.08%	3.63%	−7.94%	−1.07%	0.00%	1.11%	9.05%
SPREAD	−0.02%	3.19%	−7.41%	−0.86%	0.00%	0.82%	7.69%
PDIFF	−0.0116	0.1840	−1.0171	−0.0078	0.0019	0.0153	0.3192
ADR and Home-Country (Abnormal) Trading Volume							
AVOL_ADR	1.0259	0.3229	0.2322	0.9043	1.0193	1.1449	1.8309
AVOL_HOME	1.0185	0.3127	0.2876	0.9417	1.0165	1.0910	1.7819
AVOL_SPREAD	0.0074	0.3031	−0.8992	−0.1108	0.0061	0.1335	0.8410
Market Returns and (Abnormal) Market Trading Volume							
US_MKT	0.11%	0.89%	−2.20%	−0.32%	0.16%	0.61%	2.19%
HOME_MKT	0.09%	1.03%	−2.87%	−0.39%	0.12%	0.58%	2.55%
FX	0.02%	0.60%	−1.29%	−0.17%	0.00%	0.21%	1.31%
US_MKT_AVOL	1.0074	0.0113	0.9771	1.0009	1.0084	1.0156	1.0265
HOME_MKT_AVOL	1.0157	0.0449	0.9067	0.9931	1.0153	1.0374	1.1563
Firm Characteristics							
DIVYLD	1.64%	2.04%	0.00%	0.00%	0.97%	2.58%	8.22%
YLD_DUM	0.6605						
APPR	15.90%	44.86%	−73.19%	−11.13%	12.27%	36.18%	199.51%
APPR_DUM	0.6695						

(*continued on next page*)

TABLE 1 (continued)

---

[a] The sample is based on all foreign firms with a U.S. exchange listing for which sufficient ADR price data on CRSP and home-country price data on Datastream exist during the two years surrounding the 1997 reduction in U.S. capital gains taxes. We require that firms are listed in the U.S. for at least one year at the time of the event (May 2, 1997), have a minimum of 100 trading days (i.e., positive trading volume) in the sample period, firms' stocks actively trade during the event, and same day ADR and home-country prices are available. The table reports the number of individual firms and firm-day observations by country. We further distinguish between the following two subsamples: (1) dividend-paying firms paid dividends in the year prior to the event, and (2) appreciated firms experienced a positive stock price appreciation in the year leading to the event.

[b] The sample comprises 120,094 firm-day observations from all foreign firms with a U.S. exchange listing for which sufficient ADR price data on CRSP and home-country price data on Datastream exist during the two years surrounding the 1997 reduction in U.S. capital gains taxes.

RET_ADR and *RET_HOME* are daily stock returns (based on closing prices) for a firm's ADR in the U.S. and the underlying stock in the home country. *SPREAD* is the difference between *RET_ADR* and *RET_HOME* on the same day. The price difference, *PDIFF*, is the natural log of the ratio of the same day ADR price to the underlying stock price in the home country (adjusted for the ADR bundling ratio and translated into U.S.$). *AVOL_ADR* (*AVOL_HOME*) is the abnormal trading volume for a firm's ADR in the U.S. (the underlying stock in the home country). *AVOL_SPREAD* is the same day difference between the two measures. For each firm we calculate the abnormal trading volume by dividing each day's trading volume by the average daily trading volume in the year leading to the event. We transform trading volume data (denominated in local currencies) using natural logs. *US_MKT* and *HOME_MKT* are the daily market returns (based on closing values) for the CRSP value-weighted market index and the respective home-country market indices. *FX* is the daily currency return computed as the price relative of foreign exchange rates (foreign currency to U.S.$) minus 1. *US_MKT_AVOL* and *HOME_MKT_AVOL* represent the aggregate daily abnormal trading volume for the CRSP market index and the respective home-country market indices. *DIVYLD* is the dividend yield calculated as last fiscal year's dividend divided by stock price. *YLD_DUM* is a binary indicator variable set to 1 if a firm paid dividends in the year prior to the event. *APPR* is the buy and hold return over the prior year. *APPR_DUM* is a binary indicator variable set to 1 if a firm experienced a positive stock price appreciation in the year leading to the event.

market efficiency implies that extreme observations form part of the effect researchers are trying to capture (Kothari et al. 2005), we refrain from any data trimming in our main analyses.

Market and Currency Controls

In our returns analyses, we control for three market forces. *US_MKT* is the daily CRSP value-weighted market return. *HOME_MKT* and *FX* are the daily changes in the relevant home-country market indices and the foreign exchange rates (foreign currency relative to U.S.$), respectively. In the volume analyses, we replace the market returns with aggregate volume measures. *US_MKT_AVOL* and *HOME_MKT_AVOL*, each calculated in the same manner as the firm-level metrics, represent the daily abnormal trading volume for the CRSP market index and the respective home-country market indices. Non-U.S. data and exchange rates are collected from Datastream. Table 1, Panel B, provides descriptive statistics.

Specification of Empirical Models

To empirically test H1a and H1b we build on Lang and Shackelford (2000), who interact a proxy for a firm's sensitivity to the tax rate reduction with an event indicator and include the U.S. market return as a control. We use dividend yield in place of a dichotomous indicator to better capture the variation across the dividend-paying firms in our sample. Next, we include exchange rate fluctuations (*FX*) and home-country market returns (*HOME_MKT*) to account for the fact that home-country returns are expressed in local currencies and that price discovery is influenced by information in ADRs' countries of origin (Froot and Dabora 1999; Kim et al. 2000; Patro 2000; De Jong et al. 2005; Grammig et al. 2005). Consistent with Gagnon and Karolyi (2004), we also include lagged *PDIFF* to control for mean reversion. Finally, we include country-fixed effects to account for systematic country-level biases (e.g., time zone differences, transaction costs, and legal and institutional differences). This leads to the following empirical model for tax capitalization where $EVENT_C$ represents the five event days:

$$
\begin{aligned}
\text{Dependent Variable}_{i,t} = {} & \beta_0 + \beta_1 DIVYLD_{i,t} + \beta_2 EVENT_{Ct} \\
& + \beta_3 EVENT_{Ct} * DIVYLD_{i,t} + \beta_4 US\_MKT_t \\
& + \beta_5 HOME\_MKT_{i,t} + \beta_6 FX_{i,t} + \beta_7 PDIFF_{i,t-1} \\
& + \Sigma\beta_j \text{Country Controls}_i + \varepsilon_{i,t}. \qquad (1)
\end{aligned}
$$

Similar to Cook (2007) and Dai et al. (2008), we include past capital appreciation and its interaction with the event indicator, $EVENT_L$ (equal to 1 during the five event days) in our lock-in model. This leads to the following empirical specification to test H2a:

$$
\begin{aligned}
\text{Dependent Variable}_{i,t} = {} & \partial_0 + \partial_1 APPR_{i,t} + \partial_2 EVENT_{Lt} \\
& + \partial_3 EVENT_{Lt} * APPR_{i,t} + \partial_4 US\_MKT_t \\
& + \partial_5 HOME\_MKT_{i,t} + \partial_6 FX_{i,t} + \partial_7 PDIFF_{i,t-1} \\
& + \Sigma\partial_j \text{Country Countrols}_i + \upsilon_{i,t}. \qquad (2)
\end{aligned}
$$

For H2b we use the lock-in model except that we replace controls for market index returns with daily changes in aggregate abnormal market volume and eliminate lagged price differences:

$$\begin{aligned}\text{Dependent Variable}_{i,t} = {} & \gamma_0 + \gamma_1 APPR_{i,t} + \gamma_2 EVENT_{Lt} \\ & + \gamma_3 EVENT_{Lit} * APPR_{i,t} + \gamma_4 US\_MKT\_AVOL_t \\ & + \gamma_5 HOME\_MKT\_AVOL_{i,t} + \gamma_6 FX_{i,t} \\ & + \Sigma\gamma_j \text{Country Controls}_i + \mu_{i,t}. \end{aligned} \tag{3}$$

In the return models, we define the dependent variables as *RET_ADR*, *RET_HOME*, or the difference between the two (*SPREAD*). In the volume model, we use *AVOL_ADR*, *AVOL_HOME*, or *AVOL_SPREAD* as dependent variables; *i* and *t* represent firm and time indicators, respectively. We estimate all models using OLS with heteroscedasticity-corrected standard errors that are clustered by firm to control for within-firm correlation among variables.

We next briefly discuss the expected signs for the control variables. We expect *RET_ADR* to be positively related to both market returns. For *RET_HOME*, the role of the U.S. market is less obvious, which leads us to only predict a positive correlation with *HOME_MKT*.[21] This translates into a positive coefficient on *US_MKT* and a negative on *HOME_MKT* in the *SPREAD* regressions, consistent with the relative importance of the U.S. market differing across the two securities (e.g., Kim et al. 2000; Eun and Sabherwal 2003). Similarly, we anticipate a positive or indefinite relation between a firm's abnormal volume (i.e., *AVOL_ADR* and *AVOL_HOME*) and that of the two markets, resulting in a positive (negative) association with the U.S. (home) market in the spread regression. Prior studies suggest a negative association between exchange rates and *RET_ADR*, as well as a positive (or insignificant) relation with *RET_HOME* (e.g., Froot and Dabora 1999; Kim et al. 2000; Grammig et al. 2005). Together these effects indicate a negative relation between *FX* and *SPREAD*. We do not form any predictions on the foreign currency variable in the volume specification. Finally, if prices tend to revert toward parity, then we should find a negative (positive) coefficient on *PDIFF* using *RET_ADR* (*RET_HOME*) as dependent variable, leading to a negative effect on *SPREAD*.

IV. TESTS OF TAX CAPITALIZATION

Univariate Returns Analysis in the Capitalization Week

We begin our empirical investigation with a univariate comparison of *RET_ADR* and *RET_HOME* across dividend payout policy. Table 2 illustrates that the event period returns are statistically higher than non-event returns, regardless of firm payout policy or market. It appears that the capitalization week was a period of generally good news to the markets. ADR returns for non-dividend-payers exhibit the highest value. In addition, non-dividend-paying ADRs significantly outperform dividend-payers by 34 basis points per day during the event period as compared to an underperformance of three basis points during the non-event period. As there is no difference in mean *RET_HOME* associated with payout policy,

---

[21] Univariate analyses show that the Pearson correlation between *RET_ADR* and *RET_HOME* is only about 0.60, suggesting that market and institutional forces affect the two securities differently. Furthermore, *RET_ADR* is correlated similarly with both *US_MKT* and *HOME_MKT*, whereas *RET_HOME* has a much higher correlation with *HOME_MKT*.

TABLE 2
Summary Returns Statistics for the Tax Capitalization Week (April 30 to May 6, 1997)

	Variables	Predictions under Tax Capitalization H1a (H1b)	Event Period (1)	Non-Event Period (2)	Difference (1) − (2)
Dividend-Paying Firms	*RET_ADR*		0.82%	0.06%	0.76%***
(a)	*RET_HOME*		0.84%	0.09%	0.75%***
	SPREAD		−0.02%	−0.03%	0.01%
	n		711	78,611	
Non-Dividend-Paying Firms	*RET_ADR*		1.16%	0.03%	1.13%***
(b)	*RET_HOME*		0.86%	0.05%	0.81%***
	SPREAD		0.30%	−0.02%	0.32%
	n		414	40,358	
Difference (a) − (b)	*RET_ADR*	− (−)	−0.34%*	0.03%*	
	RET_HOME	0 (−)	−0.02%	0.04%*	
	SPREAD	− (0)	−0.32%**	−0.01%	

*, **, *** Indicate statistical significance at the 10 percent, 5 percent, and 1 percent levels (two-tailed), respectively, based on a t-test.

The sample comprises 120,094 firm-day observations from all foreign firms with a U.S. exchange listing for which sufficient ADR price data on CRSP and home-country price data on Datastream exist during the two years surrounding the 1997 reduction in U.S. capital gains taxes.

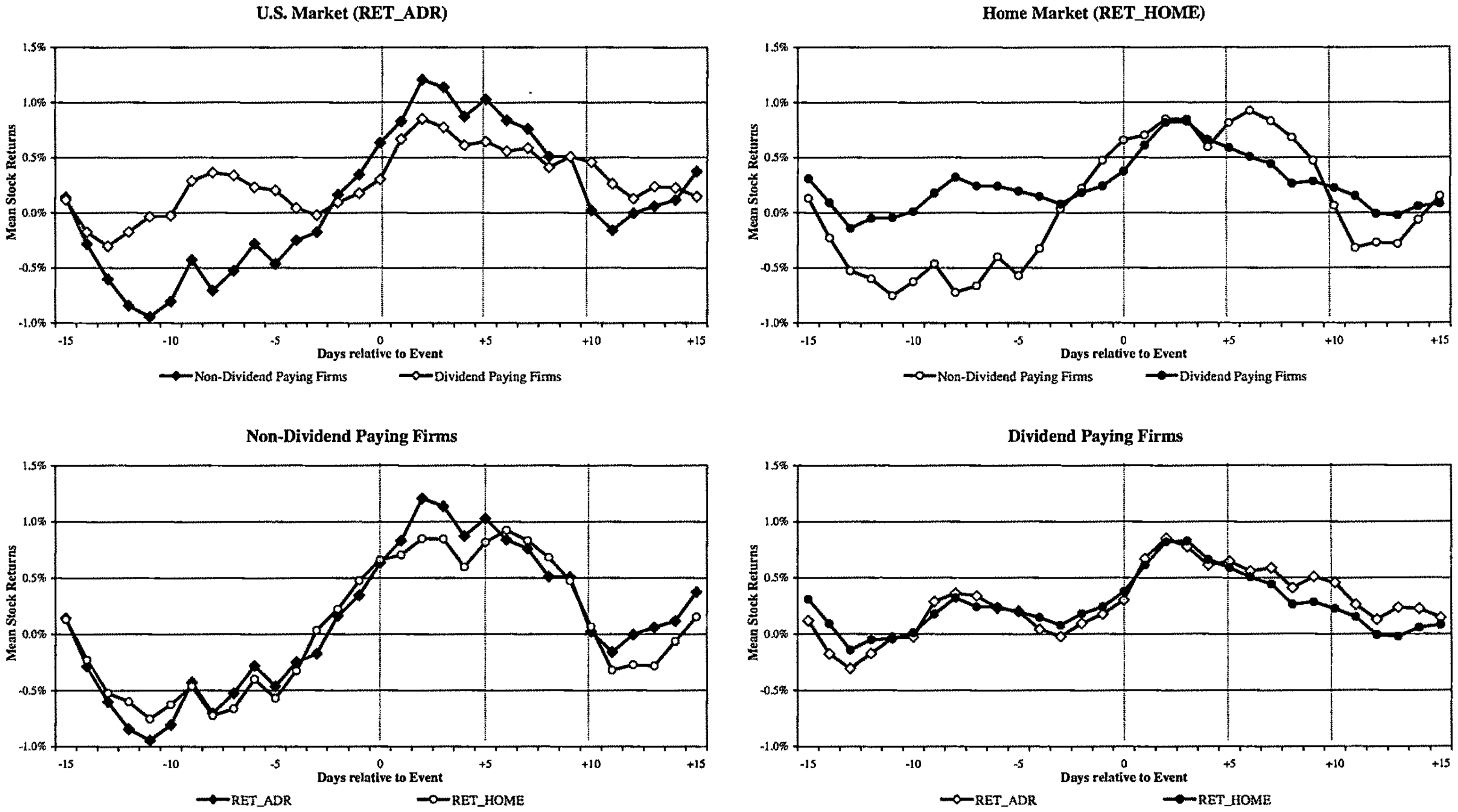
RET_ADR and *RET_HOME* are daily stock returns (based on closing prices) for a firm's ADR in the U.S. and the underlying stock in the home country. *SPREAD* is the difference between *RET_ADR* and *RET_HOME* on the same day. We distinguish between firms that did (did not) pay dividends in the year prior to the event. The table reports mean values.

a significant 32-basis-point daily *SPREAD* is created during the capitalization week. The pattern in return spreads is consistent with market prices of firms with the highest expected future capital appreciation benefiting the most from a reduction in capital gains taxes, while home-country prices seem unaffected.

To gauge the economic importance of the event, we assess whether the spread exceeds the costs of arbitrage (see also Equations (A8) and (A9) in the Appendix). In an *ad hoc* manner, we compare the 160 basis points event week spread (= 5 × 32 basis points) to the 160 basis points two-way costs of arbitrage in ADRs for institutional investors (see Chakravarty et al. 2004, Table 5) plus an average conversion fee of ten basis points (estimated by taking a nickel conversion cost as reported in Mehta [2003], deflated by the average ADR share price in our sample). Therefore, we surmise that, on average, the tax-induced pricing spread persists because the potential net gains from arbitrage are small and do not compensate adequately for risk (e.g., currency or liquidity risk).

Figure 1 pictorially confirms the above findings. Rolling five-day averages of *RET_ADR* and *RET_HOME* for dividend and non-dividend firms are presented for the 30 days surrounding the event (upper two quadrants of the graph). We observe a relative upward shift in *RET_ADR* and, less pronounced, *RET_HOME* for non-dividend-paying firms. The differential returns become more obvious in the bottom quadrants, where we compare *RET_ADR* with *RET_HOME* across dividend-paying status. While there is a clear gap of up to 36 basis points immediately following the event for non-dividend-payers, no such spread is apparent for dividend-paying firms. In fact, closer inspection reveals that the

FIGURE 1
ADR and Home-Country Returns for Dividend and Non-Dividend-Paying Firms
(Rolling Five-Day Averages, 30 Days around Event)



P 32758

spread for the non-dividend-paying firms is in the top 2 percent of the distribution of five-day spreads over the entire sample period.

Regression Analysis of Returns in the Capitalization Week

Table 3 presents coefficient estimates from estimating the tax-capitalization model. Results show that ADRs from firms with no- or low-dividend-yields significantly

TABLE 3
Returns Analysis during the Tax Capitalization Week

$$\text{Dependent Variable}_{i,t} = \beta_0 + \beta_1 DIVYLD_{i,t} + \beta_2 EVENT_{Ct} + \beta_3 EVENT_{Ct} * DIVYLD_{i,t} + \beta_4 US\_MKT_t + \beta_5 HOME\_MKT_{i,t} + \beta_6 FX_{i,t} + \beta_7 PDIFF_{i,t-1} + \Sigma\beta_j \text{Country Controls}_i + \varepsilon_{i,t}$$

	Tax Capitalization Model					
Independent Variables	**Predicted Sign H1a/1b**	***RET_ADR***	**Predicted Sign H1a (H1b)**	***RET_HOME***	**Predicted Sign H1a (H1b)**	***SPREAD***
Intercept		−0.03**		−0.01		−0.02
		(−2.50)		(−0.53)		(−1.47)
DIVYLD		0.03		0.53		−0.50
		(0.06)		(0.69)		(−0.70)
$EVENT_C$		0.23*		0.27*		−0.03
		(1.68)		(1.95)		(−0.51)
$EVENT_C * DIVYLD$	−	−9.28**	0 (−)	−2.51	− (0)	−6.76***
		(−2.47)		(−0.70)		(−3.07)
US_MKT	+	31.87***		−1.92	+	33.78***
		(13.07)		(−0.82)		(14.12)
HOME_MKT	+	67.03***	+	99.55***	−	−32.52***
		(24.66)		(27.64)		(−12.26)
FX	−	−51.65***	+	8.20*	−	−59.84***
		(−11.90)		(1.68)		(−17.82)
PDIFF	−	−0.55**	+	0.69**	−	−1.24**
		(−2.56)		(2.03)		(−2.43)
Country Controls		Yes		Yes		Yes
Number of Firms (Observations)				266 (120,094)		
R^2		7.77%		8.00%		2.79%

*, **, *** Indicate statistical significance at the 10 percent, 5 percent, and 1 percent levels (two-tailed), respectively.

The sample comprises 120,094 firm-day observations from foreign firms with a U.S. exchange listing during the two years surrounding the 1997 reduction in U.S. capital gains taxes.

We use the following three dependent variables: *RET_ADR* is the daily stock return (based on closing prices) for a firm's ADR in the U.S. *RET_HOME* is the daily stock return (based on closing prices) for a firm's underlying stock in the home country. *SPREAD* is the difference between *RET_ADR* in the U.S. and *RET_HOME* in the home country on the same day.

$EVENT_C$ is a binary indicator variable set to 1 if the observation date is on or between April 30 through May 6, 1997. See Table 1, Panel B, for a description of the remaining independent variables.

We indicate predictions based on the tax capitalization hypothesis where applicable. Country indicators are included in the regressions but not reported. The table reports OLS coefficient estimates and, in parentheses, t-statistics based on heteroscedasticity-corrected standard errors that are clustered by firm. For expositional purposes, we multiply all coefficients by 100.

outperform high-dividend-yield firms by approximately 1.20 percent during the five trading days surrounding the announcement of the tax rate cut.[22] No such response is apparent for *RET_HOME*, as the coefficient on $EVENT_C$ * *DIVYLD* is statistically insignificant. Together, this creates a negative and significant coefficient on the interaction term in the *SPREAD* specification. Moving from the 25th to the 75th percentile of the dividend yield distribution suggests a 0.87 percent increase in *SPREAD* over the event week.[23] These findings suggest that the ADR response is consistent with tax capitalization and the existence of barriers to cross-border arbitrage.

The remaining variables also reveal some interesting details. First, the significant coefficient on $EVENT_C$ in the two returns specifications suggests a generally positive reaction of our sample firms to the event even after controlling for general market conditions.[24] Next, all of the market controls as well as the lagged price differential behave as predicted and are generally statistically significant. Both *US_MKT* and *HOME_MKT* explain significant variation in *RET_ADR*. However, for *RET_HOME*, the coefficient on *HOME_MKT* is close to 1 and significant, whereas *US_MKT* is insignificant. In addition, the coefficient on *FX* is significantly negative in the *RET_ADR* specification and positive, but only marginally significant when using *RET_HOME* as the dependent variable. These differential associations are captured in the *SPREAD* regression, where we find significant relations with currency fluctuations, and U.S. and local market returns. Finally, *PDIFF* behaves consistent with prices reverting toward parity over time.[25]

Robustness Tests

To assess the sensitivity of our tax capitalization findings, we conduct a series of robustness tests and report the abbreviated results (i.e., the coefficient estimates on $EVENT_C$ and $EVENT_C$ * *DIVYLD*) in Table 4. We start by varying the sample and event periods. When we shorten the sampling period around the event to ±3 months, there is a positive and even stronger stock price reaction to the event that is decreasing in dividend yield in the U.S. but not in the home-country market, creating a significant return spread. When we reduce the event period to a three-day window (i.e., May 2 through May 6), the *RET_ADR* results are slightly attenuated, consistent with the shorter event window missing part of the capital market's response. However, there is no reaction in the home-country market resulting in a significant spread.

Next, we estimate the models without controlling for U.S. market returns. Tax capitalization predicts that any security anticipating future appreciation benefits from a capital gains tax rate cut, suggesting that $EVENT_C$ should be positive. When including *US_MKT*

---

[22] We assess the magnitude of the effect by computing the estimated change in ADR returns based on the interquartile range of the *DIVYLD* variable, i.e., (2.58 percent − 0 percent) * −9.28 percent * 5 days = −1.20 percent.

[23] An alternative explanation for our results is that the macroeconomic shock created by the reduction in the capital gains tax rates differentially alters the price formation process of non-dividend and dividend-paying firms. We run *CUSUM* analyses in order to determine whether there was a structural shift in our model. Untabulated results suggest that our regression model remained stable over the entire sample period.

[24] In untabulated analyses, we also test whether the event generated any real economic effects. We regress foreign currency, local, and U.S. market returns (our proxies for macroeconomic effects) on the event indicator, lead and lag values of the dependent variable, and the pair of market and/or currency controls not used on the left-hand side. We find no economic effect of the event on local market returns, a significant and negative effect on foreign currency returns and, as expected, a significantly positive effect on market returns in the U.S.

[25] Note that we do not control for differing firm characteristics (i.e., size, financial leverage, return on assets) since our research design assures that (in the spread regressions) each company serves as its own control. Moreover, if we include firm-fixed effects instead of country controls, the results also hold, albeit at slightly weaker levels of statistical significance.

TABLE 4
Sensitivity Analyses of the Tax Capitalization Effect

		Dependent Variables		
Independent Variables	**n**	***RET_ADR***	***RET_HOME***	***SPREAD***
(1) Reduced Sample Period (+/− 3 months)				
$EVENT_C$	30,165	0.44***	0.41***	0.03
		(3.05)	(2.93)	(0.40)
$EVENT_C * DIVYLD$		−10.33***	−4.71	−5.62***
		(−2.62)	(−1.29)	(−2.61)
(2) Reduced Event Window (3 days)				
$EVENT_C$	120,094	0.25	0.20	0.04
		(1.48)	(1.27)	(0.44)
$EVENT_C * DIVYLD$		−8.68*	0.18	−8.86**
		(−1.77)	(0.04)	(−2.52)
(3) No Control for U.S. Market Returns				
$EVENT_C$	120,094	0.47***	0.25*	0.21***
		(3.38)	(1.84)	(3.29)
$EVENT_C * DIVYLD$		−8.75**	−2.55	−6.20***
		(−2.32)	(−0.71)	(−2.83)
(4) Dividend-Paying Firms Only				
$EVENT_C$	79,322	0.09	0.27**	−0.18***
		(0.71)	(2.33)	(−3.69)
$EVENT_C * DIVYLD$		−6.91**	−3.19	−3.72**
		(−2.15)	(−1.13)	(−2.04)
(5) Canadian Firms Excluded				
$EVENT_C$	74,253	−0.02	0.35**	−0.37***
		(−0.13)	(2.60)	(−5.68)
$EVENT_C * DIVYLD$		−7.18**	−5.34	−1.83
		(−1.99)	(−1.63)	(−0.89)
(6) Winsorized Returns				
$EVENT_C$	120,094	0.22	0.28**	−0.06
		(1.56)	(2.02)	(−0.84)
$\mathrm{EVENT}_C * DIVYLD$		−9.12**	−2.95	−6.21***
		(−2.40)	(−0.81)	(−2.90)
(7) Panel-Corrected Standard Errors				
$EVENT_C$	116,233	0.22***	0.24	−0.00
		(3.82)	(0.94)	(−0.01)
$EVENT_C * DIVYLD$		−8.68	−1.39	−6.55**
		(−1.52)	(−0.14)	(−2.06)
(8) Dividend Indicator instead of Dividend Yield				
$EVENT_C$	120,094	0.29	0.25	0.04
		(1.39)	(1.19)	(0.39)
$EVENT_C * YLD\_DUM$		−0.33	−0.04	−0.29***
		(−1.49)	(−0.17)	(−2.68)

---

*, **, *** Indicate statistical significance at the 10 percent, 5 percent, and 1 percent levels (two-tailed), respectively.

The sample comprises a maximum of 120,094 firm-day observations from foreign firms with a U.S. exchange listing during the two years surrounding the 1997 reduction in U.S. capital gains taxes. The dependent variables are *RET_ADR*, *RET_HOME*, and *SPREAD*. See Table 1, Panel B, for variable definitions.

We report results for the following specifications: (1) we reduce the sample period to include the six months surrounding the 1997 reduction in capital gains taxes, (2) we shorten the event window to begin on the event day, May 2, through May 6 (3 trading days), (3) we omit U.S. market returns (*US_MKT*) from our set of control variables, (4) we limit the sample to the subset of dividend-paying firms, (5) we exclude the largest

(*continued on next page*)

TABLE 4 (continued)

sample country (i.e., Canada) from the analysis, (6) we winsorize the return metrics at the upper and lower .025 percentile, (7) we estimate Prais-Winsten regressions with panel-corrected standard errors (Beck and Katz 1995), and (8) we replace the continuous dividend yield with a binary indicator set equal to 1 if a firm paid dividends in the year prior to the event (*YLD_DUM*). The table reports only the event indicator ($EVENT_C$) and the interaction effect with dividend yield ($EVENT_C$ * *DIVYLD*), but the full set of controls is included. See Table 3 for details.

Unless stated otherwise, the table reports OLS coefficient estimates and, in parentheses, t-statistics based on heteroscedasticity-corrected standard errors that are clustered by firm. For expositional purposes, we multiply all coefficients by 100.

as a control, $EVENT_C$ captures only the average event-week return for non-dividend-paying firms *incremental* to that of the U.S. market as a whole. Thus, the inclusion of US_MKT could attenuate the coefficient on $EVENT_C$ (yet, does not affect $EVENT_C$ * *DIVYLD*, our main variable of interest).[26] The results in Table 4 confirm this conjecture. After excluding *US_MKT*, we obtain significantly positive coefficients on $EVENT_C$ in the *RET_ADR* and *SPREAD* regressions.

We then address the issue whether Canadian firms unduly influence our results. First, as Canadian firms comprise more than 70 percent of the non-dividend sample, we re-estimate the tax-capitalization model after excluding all non-dividend-payers (specification 4). Results are similar to our main analysis. Second, when we explicitly exclude Canadian firms (specification 5), we find a significantly negative coefficient on $EVENT_C$ * *DIVYLD* for ADR firms. The returns reaction in the home country is negative and marginally significant (two-tailed $p = 0.12$). Together, this creates a negative, but insignificant coefficient on the interaction term in the *SPREAD* regression. This relatively weak finding is likely due to a loss of power and our rather conservative model specification. In additional tests, not tabulated, $EVENT_C$ * *DIVYLD* is statistically significant once we exclude U.S. market returns and/or foreign exchange returns from the list of control variables.

The results are also robust to winsorizing daily returns at the top and bottom 0.025 percentile. Further, we estimate Prais-Winsten regressions to control for heteroscedasticity, within-panel serial correlation, and cross-sectional dependence, thereby producing standard errors that are more conservative than under firm-level clustering (Beck and Katz 1995). Accordingly, the *RET_ADR* results are only marginally significant (two-tailed $p = 0.13$). Nonetheless, the interaction term in the *SPREAD* specification remains significantly negative.

Finally, we replace *DIVYLD* with the binary indicator *YLD_DUM*. This analysis yields a coefficient on the interaction term that is no longer significant in the *RET_ADR* model (p-value of 14 percent, two-tailed).[27] In the *SPREAD* specification, though, the results are still consistent with tax capitalization.

We estimate several alternate model specifications but do not tabulate the results. First, we control for the possibility that dividend yield is correlated with market risk and include firm-specific betas for the ADRs and the underlying securities in the regression models.[28] None of the beta factors is statistically significant, and our inferences hold. Second, we estimate a model where we include one-day lead and lag values of the market factors (*HOME_MKT*, *US_MKT*, and *FX*). This allows us to control for asynchronous trading between the U.S. and home-country markets and to capture slow information diffusion across markets (Gagnon and Karolyi 2004). Once again, our inferences hold. Third, we identify two other periods of extreme market run-ups (i.e., October 30, 1997, and January 15, 1998) to see whether dividend yield appears to act as a proxy for some omitted market-risk factor.[29] When we rerun our analysis around these alternative extreme positive return periods, we find no evidence consistent with our tax predictions. Fourth, we estimate the models without *PDIFF* to ensure that its inclusion does not cause an estimation bias due to returns mechanically causing changes in current prices. Fifth, we include the lock-in variables ($EVENT_L$, *APPR*, and $EVENT_L$ * *APPR*) as additional controls in the tax-capitalization model. Finally, we rerun all analyses using home-country returns translated into U.S. dollars. None of these latter three specifications materially affects our results.

---

[26] Lang and Shackelford (2000) find evidence that non-dividend-paying firms in the U.S. had a 12 percent return during the week the rate cut was announced.

[27] In comparison, Lang and Shackelford (2000) report a significant 4.25 percent event-week return difference between dividend- and non-dividend-paying firms for their sample consisting of the 2,000 largest U.S. corporations. When we limit our sample to the larger, more liquid firms (i.e., those that trade every day), we find an event-week return differential using $EVENT_C$ * *YLD_DUM* of approximately 4 percent, which is significant at the 1 percent level.

V. TESTS OF CROSS-BORDER ARBITRAGE

Impact of Barriers to Arbitrage on Tax Capitalization

In this section, we examine whether cross-sectional variation in sample firms' home-country stock price reactions to the U.S. tax event can be explained by costs of arbitrage. Recall that tax capitalization implies a permanent upward shift in equilibrium prices, thereby lowering firms' costs of capital. Therefore, conditional on our evidence of a positive ADR stock price reaction, home-country stocks should mirror ADR prices if the benefits of arbitrage outweigh its costs (see the Appendix for a proper definition of the arbitrage conditions). Empirically, we split the sample into low and high costs of arbitrage groups using four proxies of firm-level liquidity and investor sophistication and report the results of the main variables of interest in Table 5.[30]

Our first proxy is the average daily U.S.$ trading volume for a firm's ADR. Trading volume captures investors' willingness to trade in the security and, *ceteris paribus*, should be inversely related to the existence of information asymmetries (Leuz and Verrecchia 2000). We partition our sample based on the sample median. Our second proxy is daily trading activity in the U.S. We define the low (high) costs of arbitrage subsample as those ADRs with (without) consecutive trading over the two-year sample period (505 days). Our third proxy measures the price impact of the order flow, which increases in illiquidity. We calculate this metric by averaging the daily Amihud (2002) illiquidity ratios for each firm in each market over the entire sample period. We then assign firms to the low costs of

---

[28] Market betas of the ADR and the underlying stock are estimated based on 100-day, one-factor market model regressions on either the U.S. stock market or the stock market in the country of the underlying security, leading to a total of four different beta factors (i.e., two for the ADR and two for the underlying security).

[29] We determine alternate dates of extreme market activity by cumulating five-day windows of CRSP's value-weighted returns. Note that the top period of market performance is our tax event, with the five days surrounding October 30, 1997 (January 15, 1998) as the second (third) strongest week during the entire sample period.

[30] Alternatively, one could measure barriers to arbitrage on the country level. However, with regard to the rather limited number of individual firms in our sample, each country-level partition heavily weighs on the cross-sectional variation in the effect we are trying to capture. For instance, when we partition our sample based on the mean equity trading costs per country (Domowitz et al. 2001), the results are generally consistent with the arbitrage story, but statistically not significant. In a similar vein, to ensure that our cross-sectional analysis does not capture the tax attributes of the average investor, we collect country-level individual capital gains tax rates to estimate $(1-t_{HC})/(1-t_{US})$, where t stands for the tax rate in the home country and the U.S., respectively. We then correlate this tax ratio (or its absolute value) with our proxies of the costs of arbitrage, but do not find evidence of a positive association.

TABLE 5
Cross-Sectional Analysis of the Impact of the Barriers to Arbitrage on the Tax Capitalization Effect

$$\text{Dependent Variable}_{i,t} = \beta_0 + \beta_1 DIVYLD_{i,t} + \beta_2 EVENT_{Ct} + \beta_3 EVENT_{Ct} * DIVYLD_{i,t} + \beta_4 US\_MKT_t + \beta_5 HOME\_MKT_{i,t} + \beta_6 FX_{i,t} + \beta_7 PDIFF_{i,t-1} + \Sigma\beta_j \text{Country Controls}_i + \varepsilon_{i,t}$$

	Average U.S.$ Trading Volume for a Firm's ADR			Number of Trading Days of a Firm's ADR		
Variables	***RET_ADR***	***RET_HOME***	***SPREAD***	***RET_ADR***	***RET_HOME***	***SPREAD***
Low Costs of Arbitrage Subsample	Equal to or Above 50th Percentile			Maximum Number (505 Trading Days)		
Number of Firms (Observations)	127 (60,262)			142 (68,815)		
DIVYLD	0.70 (1.53)	0.43 (0.92)	0.26 (0.85)	0.55 (1.23)	0.60 (1.07)	−0.05 (−0.09)
$EVENT_C$	0.21 (1.33)	0.34** (2.12)	−0.14** (−2.19)	0.28* (1.69)	0.42** (2.43)	−0.14** (−2.44)
$EVENT_C$ * *DIVYLD*	−13.96*** (−2.96)	−8.98* (−1.85)	−4.97*** (−2.85)	−13.36*** (−2.85)	−8.89* (−1.89)	−4.47*** (−2.94)
High Costs of Arbitrage Subsample	Below 50th Percentile			Below Maximum Number		
Number of Firms (Observations)	139 (59,832)			124 (51,279)		
DIVYLD	−1.05 (−1.10)	1.16 (0.74)	−2.21 (−1.37)	−0.63 (−0.58)	1.75 (0.87)	−2.38 (−1.10)
$EVENT_C$	0.28 (1.38)	0.23 (1.16)	0.05 (0.50)	0.22 (1.01)	0.12 (0.59)	0.10 (0.78)
$EVENT_C$ * *DIVYLD*	−5.92 (−1.10)	2.81 (0.56)	−8.74** (−2.34)	−5.85 (−0.98)	5.24 (0.91)	−11.09** (−2.34)
Difference in $EVENT_C$ * *DIVYLD* between Subsamples (t-stat)	(1.09)	(1.72)*	(1.04)	(1.00)	(1.82)*	(1.24)

(*continued on next page*)

TABLE 5 (continued)

	Average Price Impact in U.S. and Home Country			Ownership by Institutional Investors		
Variables	***RET_ADR***	***RET_HOME***	***SPREAD***	***RET_ADR***	***RET_HOME***	***SPREAD***
Low Costs of Arbitrage Subsample	Equal to or Below 33rd Percentile in at Least One Market			Equal to or Above 75th Percentile		
Number of Firms (Observations)	120 (57,638)			66 (30,072)		
DIVYLD	0.91**	0.48	0.43	0.78	0.49	0.30
	(2.50)	(1.11)	(1.40)	(1.50)	(0.49)	(0.30)
$EVENT_C$	0.22	0.43**	−0.21***	0.63***	0.74***	−0.12
	(1.29)	(2.58)	(−3.82)	(2.94)	(3.22)	(−1.31)
$EVENT_C$ * *DIVYLD*	−13.22**	−9.53*	−3.69**	−18.77***	−14.83*	−3.94
	(−2.61)	(−1.96)	(−2.61)	(−2.68)	(−1.95)	(−1.56)
High Costs of Arbitrage Subsample	Above 33rd Percentile in Both Markets			Below 75th Percentile		
Number of Firms (Observations)	146 (62,456)			200 (90,022)		
DIVYLD	−2.08	2.58	−4.67*	−0.34	0.51	−0.85
	(−1.56)	(1.14)	(−1.71)	(−0.59)	(0.54)	(−1.07)
$EVENT_C$	0.27	0.20	0.08	0.05	0.07	−0.01
	(1.49)	(1.09)	(0.81)	(0.31)	(0.40)	(−0.13)
$EVENT_C$ * *DIVYLD*	−7.84	1.89	−9.73**	−4.95	2.80	−7.75***
	(−1.54)	(0.40)	(−2.42)	(−1.09)	(0.65)	(−2.79)
Difference in $EVENT_C$ * *DIVYLD* between Subsamples (t-stat)	(0.79)	(1.69)*	(1.42)	(1.60)	(2.07)**	(1.35)

---

*, **, *** Indicate statistical significance at the 10 percent, 5 percent, and 1 percent levels (two-tailed), respectively.

The sample comprises a maximum of 120,094 firm-day observations from foreign firms with a U.S. exchange listing during the two years surrounding the 1997 reduction in U.S. capital gains taxes. The dependent variables are *RET_ADR*, *RET_HOME*, and *SPREAD*. See Table 1, Panel B, for all variable definitions.

We present results using four alternative metrics to assign firms to the low costs of arbitrage subsamples: (1) firms where the average daily U.S.$ trading volume for their ADR is equal to or above the 50th percentile of the volume distribution, (2) firms with consecutive trading in the U.S. (505 trading days), (3) firms where the average daily price impact captured by the Amihud (2002) illiquidity measure is equal to or below the 33rd percentile in at least one market (i.e., U.S. or home country), and (4) firms where the average proportion of institutional holdings as indicated in the Spectrum Database is equal to or above the 75th percentile.

The table reports only the main and interaction effects of dividend yield (*DIVYLD*) and the event indicator ($EVENT_C$), but the full set of controls is included. See Table 3 for details.

The table reports OLS coefficient estimates and, in parentheses, t-statistics based on heteroscedasticity-corrected standard errors that are clustered by firm. It also reports t-statistics from a fully interacted model comparing the coefficients on $EVENT_C$ * *DIVYLD* across the two subsamples. For expositional purposes, we multiply all coefficients by 100.

arbitrage subsample where the U.S. and/or the home-country illiquidity metrics are equal to or below the 33rd percentile of their respective distributions. Finally, we partition the sample based on the proportion of a firm's stock that is held by institutional investors. This variable serves as not only a proxy for liquidity, but also for investor sophistication, lower information asymmetry, lower risk of expropriation by entrenched managers, and higher firm visibility. We measure institutional ownership using the 13-F data as of the prior quarter and classify firms' stocks as more easily arbitraged if their institutional holdings are equal to or greater that the 75th percentile of the distribution.[31]

For the low costs of arbitrage subsamples, Table 5 indicates that both *RET_ADR* and *RET_HOME* are significantly decreasing in dividend yield during the event, consistent with cross-border arbitrage transferring the positive ADR response to the reduction in capital gains tax rates into home-country prices. Except for the institutional holdings' partition, where the arbitrage seems complete, there remains a significant, albeit smaller spread beyond the five-day event window. This result is consistent with the forces of arbitrage mitigating any event-related price deviations from parity, but only to the point where the benefits of doing so exceed the costs of arbitrage. So, even for highly liquid firms cross-border arbitrage is not without frictions.

Turning to the firms with relatively high costs of arbitrage, the table reports that neither *RET_ADR* nor *RET_HOME* are significantly decreasing in dividend yield during the event. Nonetheless, the pricing spread is negative and significant across all four partitions, consistent with tax capitalization *and* with barriers to arbitrage impeding prices from reaching parity within five trading days. When we compare the $EVENT_C * DIVYLD$ coefficients between the low and high costs of arbitrage subsamples, we find that the reaction in the home country is indeed more negative for easily arbitraged stocks (as evidenced by the significant t-statistics in the bottom row of Table 5).

Permanence of the Tax Capitalization Effect

Having documented a significant price increase in the no- or low-dividend-yield ADRs and, whenever costs of arbitrage are sufficiently low, in the home-country securities *during* the event week, we now investigate whether this effect persists well beyond the short-term horizon of five trading days (Gagnon and Karolyi 2004). Such evidence would be consistent with investor-level taxes increasing equilibrium prices and, in turn, decreasing firms' costs of capital. We test this claim by examining cumulative abnormal returns over 45 trading days beginning in the week immediately prior to the event (i.e., on April 23, 1997, or $t = -5$). Abnormal returns are the residuals from a one-factor market model, estimated over the year leading up to the event. We then sum the abnormal returns over the respective post-event periods.[32] The analysis is limited to the home-country securities since there is no appropriate control group for the ADRs in the U.S.

Figure 2 depicts the daily average cumulative abnormal returns across cross-listing status and dividend-paying status. In the upper left panel, we focus on firms with a U.S. exchange listing. In the week prior to the event, cumulative abnormal returns for dividend- and non-dividend-paying firms hover around zero. However, at the event, consistent with

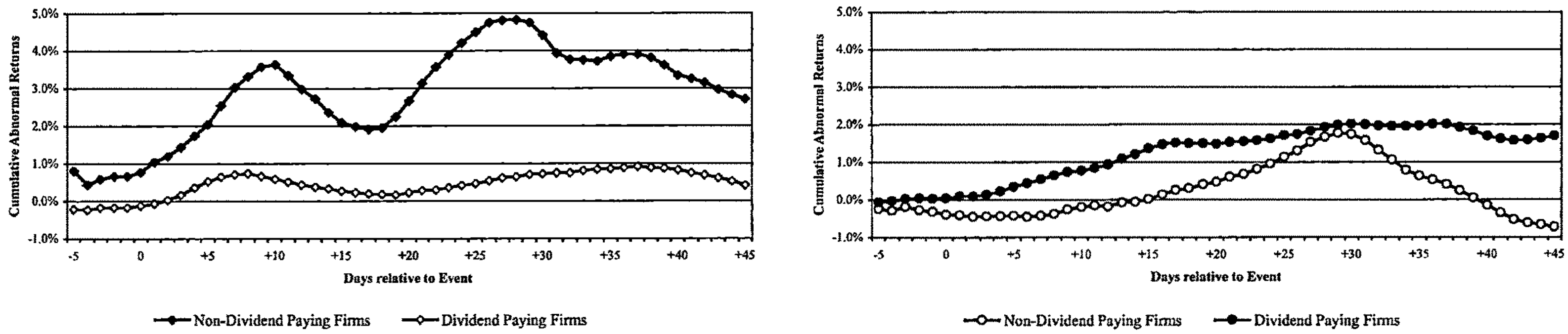
---

[31] We consider each of our firm-level proxies separately in determining its cutoff in order to account for its own unique distribution and to assure sufficient cross-sectional variation within our relatively small sample of firms. As a result, the assignment of individual firms to the low and high costs of arbitrage subsamples varies substantially across the four proxies. Nonetheless, on an aggregate level, the four partitions are highly correlated.

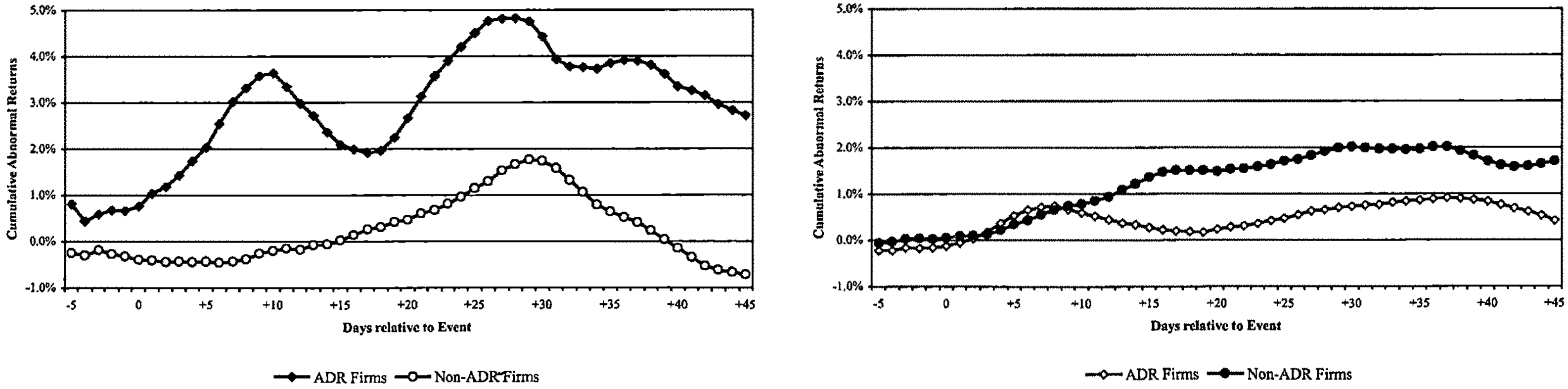
[32] The one-factor model consists of the local market index. However, results are robust to using a three-factor model comprising the U.S. market, home-country market and foreign exchange rate returns as factors. To be included in the analysis, we require a firm to have at least 30 observations in the pre-event period.

FIGURE 2
Home Market Cumulative Abnormal Returns across ADR Status and Dividend-Paying Status (Daily Averages, 50 Days Beginning Five Days before the Event Date)

ADR Firms in Home Markets **Non-ADR Firms in Home Markets**



Non-Dividend-Paying Firms in Home Markets **Dividend-Paying Firms in Home Markets**



tax capitalization, the cumulative abnormal returns of the non-dividend-paying firms increase; a result not apparent for the dividend-paying firms. This positive cumulative abnormal returns spread persists for at least two months, suggesting a permanent, rather than temporary, shift in home-country prices.

The upper right panel provides the analysis for the non-ADR sample, which serves as a benchmark because the securities without U.S. cross-listings should be relatively unaffected by U.S. tax legislation. We identify the sample of non-ADR firms in a manner similar to our ADR sample selection procedure.[33] Note that the cumulative abnormal returns of these securities are close to zero over the entire period without apparent distinction across dividend-paying status.

The abnormal returns differential between firms with a U.S. cross-listing and their home-country peers is depicted in the bottom half of Figure 2. For the non-dividend-paying firms, a return spread of about 300 basis points develops when the U.S. tax rate cut is announced and persists for at least two months after the event. No such spread is evident for the dividend-paying firms. In untabulated analyses, we combine the return spreads across dividend-paying status in a difference-in-differences design, and find an increase in the difference between the spreads developing at the event, which levels out around trading day +10. This difference-in-differences is significantly positive at the 5 percent level (two-tailed) until at least day +45. Overall, the evidence of an increase in equity values persisting beyond the event window is consistent with U.S. tax rates affecting foreign firms' costs of capital.

Tax Capitalization and Capital-Raising Activity

To further support the reduction in cost of capital conjectures, we next investigate the equity-raising activities of our sample firms in order to determine whether—as one would expect—these firms benefited more from the reduction in capital gains tax rates. To proxy for a firm's expected propensity to raise capital, we use two binary variables to split the 142 most-liquid firms (i.e., those with low costs of arbitrage) into subgroups. First, we distinguish between cross-listed firms that are allowed to engage in a U.S. public offering (Level III firms) and firms that are not (Level II firms). Second, we identify ADRs with and without actual capital-raising activities in the U.S. occurring over the five-year period after the 1997 event (as indicated in the Securities Data Company database).

Table 6 presents the coefficients of interest from estimating the tax-capitalization model across these partitions. The home security (as well as the ADR) response on $EVENT_C * DIVYLD$ is significantly negative for firms with higher probabilities of accessing U.S. capital markets in the future. The same coefficients are negative, but generally insignificant for firms that are less likely to raise capital in the U.S. Although not significant in a statistical sense, the direction of the differential response across the two partitions is consistent with firms that have a greater proclivity of raising capital benefiting more from the reduction in U.S. investor-level taxes.

---

[33] We start with selecting all non-U.S., non-ADR firms with price data on Datastream and dividend information on Worldscope. Next, we require that firms actively traded for at least one day during our tax capitalization event week. To ensure informative prices, we further eliminate firms that were not listed one year prior to the event, or with average daily trading volume of less than U.S.$50,000.

TABLE 6
Cross-Sectional Analysis of the Role of Capital Raising on the Tax Capitalization Effect

$$\text{Dependent Variable}_{i,t} = \beta_0 + \beta_1 DIVYLD_{i,t} + \beta_2 EVENT_{Ct} + \beta_3 EVENT_{Ct} * DIVYLD_{i,t} + \beta_4 US\_MKT_t + \beta_5 HOME\_MKT_{i,t} + \beta_6 FX_{i,t} + \beta_7 PDIFF_{i,t-1} + \Sigma\beta_j \text{Country Controls}_i + \varepsilon_{i,t}$$

	Exchange-Listing Status			Actual Capital Raising		
Variables	***RET_ADR***	***RET_HOME***	***SPREAD***	***RET_ADR***	***RET_HOME***	***SPREAD***
Capital Raising Subsample		Level III Firms			Issuance of Equity Capital in 1997 or Later	
Number of Firms (Observations)		65 (31,513)			32 (15,582)	
DIVYLD	0.86*	0.65	0.21	1.93	3.40**	−1.47
	(1.82)	(0.97)	(0.42)	(1.29)	(2.46)	(−1.69)
$EVENT_C$	0.33	0.59***	−0.26***	0.60**	0.68**	−0.08
	(1.46)	(2.68)	(−3.75)	(2.09)	(2.43)	(−0.84)
$EVENT_C * DIVYLD$	−19.11**	−15.39**	−3.72*	−23.46**	−18.29*	−5.16*
	(−2.50)	(−2.20)	(−1.99)	(−2.30)	(−2.03)	(−1.72)
Non-Capital Raising Subsample		Level II Firms			No Issuance of Equity Capital after 1996	
Number of Firms (Observations)		77 (37,302)			110 (53,233)	
DIVYLD	1.09	−0.44	1.52	1.07*	−0.27	1.34
	(0.58)	(−0.24)	(0.46)	(1.71)	(−0.40)	(1.41)
$EVENT_C$	0.20	0.27	−0.08	0.19	0.35*	−0.16**
	(0.81)	(1.08)	(−0.97)	(0.98)	(1.73)	(−2.38)
$EVENT_C * DIVYLD$	−7.02	−4.69	−2.33	−10.63**	−7.07	−3.56*
	(−1.05)	(−0.68)	(−1.01)	(−2.04)	(−1.33)	(−1.93)
Difference in $EVENT_C * DIVYLD$ between Subsamples (t-stat)	(0.97)	(1.27)	(1.00)	(1.05)	(1.17)	(0.10)

(continued on next page)

TABLE 6 (continued)

*, **, *** Indicate statistical significance at the 10 percent, 5 percent, and 1 percent levels (two-tailed), respectively.

The sample comprises 68,815 firm-day observations from foreign firms with a U.S. exchange listing during the two years surrounding the 1997 reduction in U.S. capital gains taxes. We limit the sample to firms with consecutive trading in the U.S. (505 trading days). The dependent variables are *RET_ADR*, *RET_HOME*, and *SPREAD*. See Table 1, Panel B, for all variable definitions.

We use two metrics to assign firms to the capital-raising (non-capital-raising) subsamples: (1) the exchange-listing status in the U.S., which allows firms to either engage in a U.S. public equity offering (Level III firms) or not (Level II firms), and (2) firms' actual equity capital-raising activity on NYSE, NASDAQ, or AMEX over the period 1997 to 2004. We use the Securities Data Company database to identify actual capital issuances as well as to determine the exchange-listing status of Canadian firms.

The table reports only the main and interaction effects of dividend yield (*DIVYLD*) and the event indicator ($EVENT_C$), but the full set of controls is included. See Table 3 for details.

The table reports OLS coefficient estimates and, in parentheses, t-statistics based on heteroscedasticity-corrected standard errors that are clustered by firm. It also reports t-statistics from a fully interacted model comparing the coefficients on $EVENT_C$ * *DIVYLD* across the two subsamples. For expositional purposes, we multiply all coefficients by 100.

VI. TESTS OF LOCK-IN BEHAVIOR

Univariate Returns and Volume Analysis in the Lock-In Week

Table 7 reports the univariate comparison of *RET_ADR* and *RET_HOME* for appreciated versus depreciated firms. Panel A provides a comparison of the mean returns between the lock-in event and non-event periods. Univariate results indicate that firms generally perform better during the week when the reduced capital gains tax rate becomes effective. It appears that the lock-in week coincided with generally good news to the markets. A comparison of mean returns across capital appreciation status illustrates that depreciated ADRs significantly outperform their appreciated counterparts (positive 46 basis points per day) during the event week as compared to the non-event period (negative 17 basis points per day). Thus, the returns pattern in the U.S. suggests a price drop in appreciated stocks consistent with the lock-in effect.

However, the response of ADRs' underlying home-country securities, which serve as a near perfect control group, reveals confounding evidence. *RET_HOME* reacts very similarly to *RET_ADR* with depreciated home-country stocks significantly outperforming (underperforming) their appreciated counterparts during the event (non-event) period (a positive 43 versus a negative 16 basis points per day). Therefore, no significant *SPREAD* is created, where lock-in theory predicts a negative spread.

In Panel B of Table 7, we report the same univariate comparisons for mean abnormal trading volume. Trading activity in appreciated stocks is always higher than in depreciated stocks and, in most cases, the difference is significant. However, we do not find evidence of differential trading volume between the event and non-event periods, regardless of firms' capital appreciation. Conversely, lock-in predicts a temporary surge in volume, due to investors' willingness to unload their previously locked-in shares.

Regression Analysis of Returns and Volume in the Lock-In Week

Table 8, Panel A, provides coefficient estimates from the lock-in effect model using returns. It shows that past capital appreciation is negatively associated with ADR returns during the lock-in week. Moving from the 25th to the 75th percentile of the *APPR* distribution, the returns differential between low versus high appreciation firms is about 166 basis points. However, the coefficient on $EVENT_L$ * *APPR* is similar in terms of magnitude and statistical significance in the *RET_HOME* specification, indicating that during the lock-in week, highly appreciated home-country stocks also underperform their home-country counterparts. Together this creates a return spread that is close to zero. This finding is not only inconsistent with lock-in behavior, but also contrary to our cross-border arbitrage predictions. Because the lock-in effect's impact on prices stems from a supply-side surge in volume, the firm's fundamentals should not be affected. In fact, arbitrageurs should drive *RET_ADR* quickly back to the original equilibrium prices and not cause any price deviations in *RET_HOME*. All the control variables behave as predicted and are highly significant. Also, the positive coefficient on $EVENT_L$ in the two returns specifications supports the univariate findings of a positive marginal event response even after taking into account general market conditions.

Panel B of Table 8 presents the tests of trading volume. Contrary to expectations under the lock-in hypothesis, the coefficient on $EVENT_L$ * *APPR* is significantly negative in the *AVOL_ADR* and *AVOL_HOME* regressions, rendering the spread between the two insignificant. The evidence suggests that the observed drop in prices for highly appreciated ADRs is accompanied by a *decrease* in trading volume during the event week. The same relation applies to the underlying home-country stocks. Overall, the evidence from our univariate and multivariate returns and volume analyses does not support lock-in behavior.

TABLE 7
Summary Returns and Volume Statistics for the Lock-In Week
(May 7 to May 13, 1997)

Panel A: ADR Stock Returns, Home-Country Stock Returns, and Return Spreads

	Variables	Predictions under Lock-In Effect H2a	Event Period (1)	Non-Event Period (2)	Difference (1) − (2)
Appreciated Firms	*RET_ADR*		0.28%	0.11%	0.17%*
(a)	*RET_HOME*		0.26%	0.13%	0.13%
	SPREAD		0.02%	−0.02%	0.04%
	n		805	79,599	
Non-Appreciated Firms	*RET_ADR*		0.74%	−0.06%	0.80%***
(b)	*RET_HOME*		0.69%	−0.03%	0.72%***
	SPREAD		0.05%	−0.03%	0.08%
	n		396	39,294	
Difference (a) − (b)	*RET_ADR*	−	−0.46%***	0.17%***	
	RET_HOME	0	−0.43%***	0.16%***	
	SPREAD	−	−0.03%	0.01%	

Panel B: Abnormal U.S. and Home-Country Trading Volume and Volume Spreads

	Variables	Predictions under Lock-In Effect H2b	Event Period (1)	Non-Event Period (2)	Difference (1) − (2)
Appreciated Firms	*AVOL_ADR*		1.045	1.037	0.008
(a)	*AVOL_HOME*		1.030	1.026	0.004
	AVOL_SPREAD		0.015	0.011	0.004
	n		805	79,599	
Non-Appreciated Firms	*AVOL_ADR*		0.991	1.003	−0.012
(b)	*AVOL_HOME*		1.008	1.004	0.004
	AVOL_SPREAD		−0.017	−0.001	−0.016
	n		396	39,294	
Difference (a) − (b)	*AVOL_ADR*	+	0.054***	0.034***	
	AVOL_HOME	0	0.022	0.022***	
	AVOL_SPREAD	+	0.032*	0.012***	

---

*, **, *** Indicate statistical significance at the 10 percent, 5 percent, and 1 percent levels (two-tailed), respectively, based on a t-test.

The sample comprises 120,094 firm-day observations from all foreign firms with a U.S. exchange listing for which sufficient ADR price data on CRSP and home-country price data on Datastream exist during the two years surrounding the 1997 reduction in U.S. capital gains taxes. *RET_ADR* and *RET_HOME* are daily stock returns (based on closing prices) for a firm's ADR in the U.S. and the underlying stock in the home country. *SPREAD* is the difference between *RET_ADR* and *RET_HOME* on the same day. *AVOL_ADR* (*AVOL_HOME*) is the abnormal trading volume for a firm's ADR in the U.S. (the underlying stock in the home country). *AVOL_SPREAD* is the same day difference between the two measures.

For each firm we calculate the abnormal trading volume by dividing each day's trading volume by the average daily trading volume in the year leading to the event. We transform trading volume data (denominated in local currencies) using natural logs. We distinguish between firms that experienced a stock price appreciation (depreciation) in the year immediately prior to the event.
The table reports mean values.

TABLE 8
Returns and Volume Analysis during the Lock-In Week

Panel A: ADR Stock Returns, Home-Country Stock Returns, and Return Spreads

$$\text{Dependent Variable}_{i,t} = \partial_0 + \partial_1 APPR_{i,t} + \partial_2 EVENT_{Lt} + \partial_3 EVENT_{Lt} * APPR_{i,t} + \partial_4 US\_MKT_t + \partial_5 HOME\_MKT_{i,t} + \partial_6 FX_{i,t} + \partial_7 PDIFF_{i,t-1} + \Sigma\partial_j \text{Country Controls}_i + \upsilon_{i,t}$$

	Lock-In Effect Model					
Independent Variables	**Predicted Sign H2a**	***RET_ADR***	**Predicted Sign H2a**	***RET_HOME***	**Predicted Sign H2a**	***SPREAD***
Intercept		−0.04***		−0.01		−0.03**
		(−3.78)		(−0.51)		(−2.18)
APPR		0.06*		0.05		0.01
		(1.95)		(1.41)		(0.22)
EVENT$_L$		0.22***		0.17**		0.05
		(2.64)		(2.03)		(1.19)
*EVENT$_L$ * APPR*	−	−0.70***	0	−0.67***	−	−0.02
		(−3.36)		(−3.07)		(−0.19)
US_MKT	+	32.00***		−1.63	+	33.63***
		(13.25)		(−0.70)		(14.13)
HOME_MKT	+	67.05***	+	99.59***	−	−32.54***
		(24.64)		(27.59)		(−12.24)
FX	−	−51.69***	+	8.08*	−	−59.78***
		(−11.89)		(1.65)		(−17.81)
PDIFF	−	−0.54**	+	0.69**	−	−1.24**
		(−2.58)		(2.00)		(−2.41)
Country Controls		Yes		Yes		Yes
Number of Firms (Observations)				266 (120,094)		
R^2		7.78%		8.00%		2.79%

(*continued on next page*)

TABLE 8 (continued)

Panel B: Abnormal U.S. and Home-Country Trading Volume and Volume Spreads

$$\text{Dependent Variable}_{i,t} = \gamma_0 + \gamma_1 APPR_{i,t} + \gamma_2 EVENT_{L,t} + \gamma_3 EVENT_{L,t} * APPR_{i,t} + \gamma_4 US\_MKT\_AVOL_t + \gamma_5 HOME\_MKT\_AVOL_{i,t} + \gamma_6 FX_{i,t} + \Sigma\gamma_j \text{Country Controls}_i + \mu_{i,t}$$

	Lock-In Effect Model					
Independent Variables	**Predicted Sign H2b**	***AVOL_ADR***	**Predicted Sign H2b**	***AVOL_HOME***	**Predicted Sign H2b**	***AVOL_SPREAD***
Intercept		−202.90*** (−6.14)		−48.75* (−1.69)		−154.15*** (−5.08)
APPR		11.13*** (7.29)		11.37*** (7.91)		−0.24 (−0.23)
$EVENT_L$		1.27 (1.04)		1.32 (1.17)		−0.05 (−0.04)
$EVENT_L$ * *APPR*	+	−5.88** (−2.12)	0	−5.31* (−1.81)	+	−0.58 (−0.24)
US_MKT_AVOL	+	269.42*** (8.24)		60.74** (2.03)	+	208.67*** (6.11)
HOME_MKT_AVOL	+	31.83*** (6.27)	+	85.95*** (11.09)	−	−54.12*** (−5.46)
FX		−1.37 (−0.12)		−8.95 (−1.45)		7.58 (0.79)
Country Controls		Yes		Yes		Yes
Number of Firms (Observations)				263 (118,697)		
R^2		4.25%		4.69%		1.44%

(continued on next page)

TABLE 8 (continued)

*, **, *** Indicate statistical significance at the 10 percent, 5 percent, and 1 percent levels (two-tailed), respectively.

The sample comprises a maximum of 120,094 firm-day observations from foreign firms with a U.S. exchange listing during the two years surrounding the 1997 reduction in U.S. capital gains taxes.

In Panel A, we use the following three dependent variables: (1) *RET_ADR* is the daily stock return (based on closing prices) for a firm's ADR in the U.S., (2) *RET_HOME* is the daily stock return (based on closing prices) for a firm's underlying stock in the home country, and (3) *SPREAD* is the difference between RET_ADR in the U.S. and *RET_HOME* in the home country on the same day.

In Panel B, the dependent variables are the abnormal trading volume for a firm's ADR in the U.S. (*AVOL_ADR*), the underlying stock in the home country (*AVOL_HOME*) and the same day difference between the two (*AVOL_SPREAD*). For each firm we calculate the abnormal trading volume by dividing each day's trading volume by the average daily trading volume in the year leading to the event. We transform trading volume data (denominated in local currencies) using natural logs. $EVENT_L$ is a binary indicator variable set to 1 if the observation date is on or between May 7 through 13, 1997. See Table 1, Panel B, for a description of the remaining independent variables.

We indicate predictions based on the lock-in effect hypothesis where applicable. Country indicators are included in the regressions but not reported. The table reports OLS coefficient estimates and, in parentheses, t-statistics based on heteroscedasticity-corrected standard errors that are clustered by firm. For expositional purposes, we multiply all coefficients by 100.

However, our results shed light on concurrent work by Cook (2007) and Dai et al. (2008), who purport to find evidence of lock-in around the same event, but using the population of U.S. publicly traded firms only. Both studies find that appreciated firms significantly underperform depreciated firms. Yet, the primary focus of Cook (2007) is on returns around the announcement week of the 1997 budget accord, which did not provide investors with information regarding the effective date of the rate reduction. Furthermore, there is no analysis of the volume response around the effective date of the event. Dai et al. (2008), on the other hand, utilize the same lock-in event period as we do, and, consistent with our findings, document negative event period returns that are decreasing in past appreciation. They fail, though, to document that their return result is a manifestation of abnormal trading volume. Our paired-securities approach clarifies this finding by illustrating that ADRs' underlying home-country securities mirror the price *and* volume reaction of the cross-listed stocks, thereby further challenging the lock-in interpretation of Dai et al. (2008).

VII. CONCLUSIONS

In this study, we examine the interplay between shareholder-level taxes and costly cross-border arbitrage on the pricing of cross-listed securities. We conduct an event study in the weeks around the May 1997 budget accord and use dividend yield and past appreciation as proxies for a security's sensitivity to the unanticipated proposal of a reduction in capital gains tax rates. We measure barriers to cross-border arbitrage using several firm-level proxies of trading activity and investor sophistication. Our paired-securities approach allows us to shed light on the relation between a tax-induced shock to equity prices with identical pre-tax cash flows, yet different after-tax cash flows, and ultimately tackle the question of whether a firm's global investor base might affect its cost of equity capital.

In our first set of tests, we find that during the tax capitalization week, based on the inter-quartile range, no- or low-dividend-yield firms outperform high-dividend-yield firms by 120 basis points. No such pattern is generally apparent for the underlying securities in the ADRs' home countries, creating tax-induced pricing spreads of up to 87 basis points over the five-day event period. This result is consistent with the marginal investor being a U.S.-taxable individual who capitalizes personal taxes into security prices.

Next, we find that ADRs' underlying home-country securities also exhibit an increase in prices, but only when costs of arbitrage are low, implying that the forces of cross-border arbitrage constrain tax-induced pricing spreads to the arbitrage tunnel. Further analyses show that our sample of cross-listed firms continues to outperform a sample of non-ADR home-country peers for at least 45 days beyond the event, suggesting that the documented price reactions reflect more than just a temporary mispricing during the event period. Moreover, ADR firms' reactions to the event are associated with their proclivity to raise new capital. Taken together, our findings lend support to the conjecture that the effects of U.S. investor-level tax legislation reverberates into international asset prices.

Finally, we find that prices for appreciated ADRs decreased during the lock-in week. However, home-country security prices decreased as well. Closer inspection of the securities' trading volume reveals that this reaction is not attributable to short-term liquidity effects from a surge in selling by investors with relatively high unrealized capital gains tax liabilities. In fact, trading volume is below normal for stocks with positive share appreciation. Overall, we do not find evidence of lock-in behavior.

Our study is subject to several limitations. First, the analysis narrowly focuses on a single tax change in the capital gains tax rate. While this provides for a relatively clean setting, it does limit our ability to interpret how U.S. personal taxes affect equity prices in a broader sense. Second, firms with cross-listings represent a special subset among the listed corporations in a country, again potentially limiting the generalizability of the results. Third, although our results indicate an interplay between tax capitalization and cross-border arbitrage in the price formation of cross-listed shares, it is not possible to completely disentangle the two competing sources. Fourth, our evidence on the impact of a change in U.S. tax laws on foreign firms' costs of capital is indirect in nature and consistent with, but far from conclusive of, such an effect. Further evidence using more direct proxies of cost of capital is needed to support such a link. Finally, as with all event studies, our results are contingent on the identification of an event that is unconfounded by other (non-tax) factors.

APPENDIX

ARBITRAGE, TAXES, AND THE PRICING OF CROSS-LISTED SECURITIES

In this appendix we develop a model describing the relation between local capital gains tax rates and the prices of identical securities that simultaneously trade in two separate economies. This allows us to analyze the role of arbitrage across the two markets, which we then apply to our specific setting, i.e., the passage of the 1997 budget accord.

General Relation between Tax Rates and Returns of Cross-Listed Securities

No Arbitrage

We start with the assumption that a representative (or average) investor in the U.S., I_{US}, determines ADR prices and a representative (or average) investor in the home country, I_{HC}, determines the underlying home-country stock prices and, other things equal, both demand an expected after-tax return of r_{US} and r_{HC}, respectively. We also assume, for simplicity, that the securities we are investigating pay no dividends, but are traded in a market where dividend-paying securities exist. Suppose that the applicable capital gains tax rate on ADRs in the U.S. is t_{US} and on home-country securities is t_{HC}. This leads to the following two *independent* relations:

$$R_{US}(1 - t_{US}) = r_{US}, \text{ and} \tag{A1}$$

$$R_{HC}(1 - t_{HC}) = r_{HC}, \tag{A2}$$

where R_{US} is the before-tax expected ADR return and R_{HC} is the corresponding before-tax expected home-country stock return. Thus, the return spread is defined by the following ratio:

$$\frac{R_{US}}{R_{HC}} = \frac{r_{US}}{r_{HC}} \cdot \frac{(1 - t_{HC})}{(1 - t_{US})}. \tag{A3}$$

Therefore, R_{US} can deviate from R_{HC} whenever $r_{US} \neq r_{HC}$ and/or $t_{US} \neq t_{HC}$, and the return spread is increasing in t_{US}. Absent arbitrage, these return differences can lead to permanent pricing spreads between ADR and home-country securities.

Complete Arbitrage

Now, let us consider the forces of arbitrage. In a world without frictions, instead of holding the ADR, I_{US} could also invest in the home-country security and, assuming that capital gains taxes are imposed by an individual's country of residence, receive an after tax return of:

$$R_{HC}(1 - t_{US}). \tag{A4}$$

As I_{US} is indifferent between either of the two securities represented by Equation (A1) and Equation (A4) (which both represent the same underlying expected cash flows), it follows that the before-tax return is to be determined concurrently across the two markets and that R_{US} is equal to R_{HC}. Thus, the before-tax return spread in Equation (A3) will be arbitraged away immediately, and there will be no resulting difference in prices between the ADR and the home-country security:[34]

$$R_{US}(1 - t_{US}) = R_{HC}(1 - t_{US}) = r_{US}. \tag{A5}$$

Under this scenario, capital gains taxes and, consequently, changes in capital gains tax rates have no impact on the return spread.

Costly Arbitrage

Yet, presumably, there is some cost to I_{US} of investing in the home-country security (e.g., currency risk, currency translation fees, transaction costs, time zone differences), c_{US}, leading us to adjust Expression (A5) as follows:

$$R_{US}(1 - t_{US}) = (R_{HC} - c_{US})(1 - t_{US}) = r_{US}. \tag{A6}$$

Note that c_{US}, modeled as a percentage, is exogenous and not determined by either t_{US} or t_{HC}. Since c_{US} effectively reduces the anticipated return on the home-country security, these costs are tax deductible. For ease of exposition, we assume that $R_{HC} > R_{US}$ and, therefore, $c_{US} > 0$. Hence, in equilibrium, the adjusted ratio of the return spread is:

$$\frac{R_{US}}{R_{HC}} = 1 - \frac{c_{US}}{R_{HC}}. \tag{A7}$$

Assuming that R_{US} represents the fundamental cost of capital, I_{US} is indifferent between holding either of the two securities, but we could observe a return spread different from parity, where Equation (A7) sets the lower bound.[35] Put differently, whenever:

$$R_{HC} - R_{US} > c_{US}, \tag{A8}$$

I_{US} will invest in the underlying home-country security, thereby driving prices to the equilibrium ratio. On the other hand, I_{US} will continue to hold the ADR (as opposed to the underlying home-country security) whenever:

$$R_{HC} - R_{US} \leq c_{US}, \tag{A9}$$

and the no-arbitrage solution of Equation (A3) results. Given these costly arbitrage conditions, pricing spreads between ADR and home-country securities can exist up to the point where returns to be gained from arbitraging the pricing spread are less than the corresponding costs of investing in the other market.

---

[34] Note that the same analysis also holds for the investor in the home country of the underlying security, I_{HC}, or a tax-exempt investor on either side, i.e., in the U.S. or the security's home country.

[35] If we posit that $R_{HC} < R_{US}$, then the cost of arbitrage, c_{US}, would assume a negative value. By allowing c_{US} to be either positive or negative, pre-tax returns in the U.S. can fall either below or above the corresponding home-country returns before investor-level taxes. In addition, allowing R_{US} to represent the fundamentals is merely for illustrative purposes. Our analysis also holds, only with a different direction of arbitrage, when we allow R_{HC} to represent the "true" cost of capital.

Effect of Changes in Tax Rates on Returns of Cross-Listed Securities

No Arbitrage

If the U.S. capital gains tax rate on equity drops from t_{US} to t'_{US}, then, according to Equation (A1), before-tax expected ADR return R_{US} decreases to R'_{US}. This causes ADR prices in the U.S. to increase, consistent with a reduction in cost of capital. At the same time, home-country prices are not affected, yielding the new equilibrium return spread defined by the new tax rates:

$$\frac{R'_{US}}{R_{HC}} = \frac{r_{US}}{r_{HC}} \cdot \frac{(1 - t_{HC})}{(1 - t'_{US})} < \frac{R_{US}}{R_{HC}}. \tag{A10}$$

Thus, an increase in ADR prices after the passage of the 1997 budget accord represents a downward revision of before-tax expected ADR returns (or cost of capital) due to the change in U.S. investor-level tax rates. The resulting ratio of return spreads reflects the lower tax rate t'_{US}, and is smaller than the old ratio.

Costly Arbitrage

Again, suppose that t_{US} is lowered to t'_{US}. Following the logic in the previous section, the expected ADR returns in the U.S. decrease from R_{US} to R'_{US}. In the presence of costly arbitrage, the tax rate cut and resulting drop to R'_{US} lead to two potential pricing spread scenarios. First, we consider the case where the difference between R_{HC} and R'_{US} falls within the arbitrage tunnel, i.e., Equation (A9) holds. Because the costs of cross-border arbitrage are too high, neither I_{US} nor I_{HC} invest in the other market's security. That is, I_{US} continues to hold the ADR, leading to the following return ratio:

$$\frac{R'_{US}}{R_{HC}} = \frac{r_{US}}{r_{HC}} \cdot \frac{(1 - t_{HC})}{(1 - t'_{US})} < \frac{R_{US}}{R_{HC}} \geq 1 - \frac{c_{US}}{R_{HC}}. \tag{A11}$$

Equation (A11) is equivalent to the no-arbitrage solution in Equation (A10), with an increase in the ADR's price (due to the lower costs of capital) and no change in the cost of capital of the home-country security. Thus, after the tax rate cut, the return ratio is smaller than before, reflecting the reduction from t_{US} to $t'_{US.}$ However, the decrease in the ratio is not large enough to push the spread outside the bounds of the arbitrage tunnel. This case corresponds to H1a.

The second scenario considers a setting with positive returns to arbitrage. Now, after the decrease in expected before-tax ADR returns (due to the lower tax rate), the difference between R_{HC} and R'_{US} satisfies the arbitrage condition defined in Equation (A8). Consequently, as $R_{HC} > R_{US}$, it follows that $R_{HC} > R'_{US}$, so it is beneficial for I_{US} to go long in the ADR's underlying home-country stock, thereby increasing the demand in the home-country security.[36] This buying pressure continues until R_{HC} sufficiently decreases to R'_{HC}, so that:

---

[36] If $R'_{US} > R_{HC}$, then arbitrage strategy would call for I_{US} to go short in the underlying home-country stock, thereby increasing the supply. The same logic applies to I_{HC} when trying to exploit the above arbitrage opportunities.

$$\frac{R'_{US}}{R'_{HC}} = 1 - \frac{c_{US}}{R'_{HC}}. \tag{A12}$$

At this point, ADR and home-country security prices have reached the new equilibrium level, and no further arbitrage occurs. The direction of the resulting change in the ratio of returns is ambiguous, as the effects of t'_{US} on R'_{US} and R'_{HC} tend to offset each other. Note that in such a scenario (leading to H1b) not only ADR prices, but also home-country stock prices (through the channel of cross-border arbitrage) react to the change in t_{US}.

REFERENCES

Amihud, Y. 2002. Illiquidity and stock returns: Cross-section and time-series effects. *Journal of Financial Markets* 5: 31–56.

Ammer, J., S. Holland, D. Smith, and F. Warnock. 2006. Look at me now: What attracts U.S. shareholders? NBER Working Paper No. W12500. Cambridge, MA: NBER.

Amoako-Adu, B., M. Rashid, and M. Stebbins. 1992. Capital gains tax and equity values: Empirical test of stock price reaction to the introduction and reduction of capital gains tax exemption. *Journal of Banking & Finance* 16: 275–287.

Ayers, B., C. Lefanowicz, and J. Robinson. 2003. Shareholder taxes in acquisition premiums: The effect of capital gains taxation. *The Journal of Finance* 58: 2783–2801.

———, ———, and ———. 2004. The effect of shareholder-level capital gains taxes on acquisition structure. *The Accounting Review* 79: 859–887.

———, ———, and ———. 2007. Capital gains taxes and acquisition activity: Evidence of the lock-in effect. *Contemporary Accounting Research* 24: 315–344.

———, O. Li, and J. Robinson. 2008. Tax induced trading around the Taxpayer Relief Act of 1997. *The Journal of the American Taxation Association* 30: 77–100.

Beck, N., and J. Katz. 1995. What to do (and not to do) with time-series cross-section data. *The American Political Science Review* 89: 634–647.

Bradshaw, M., B. Bushee, and G. Miller. 2004. Accounting choice, home bias, and U.S. investment in non-U.S. firms. *Journal of Accounting Research* 42: 795–841.

Callaghan, S., and C. Barry. 2003. Tax-induced trading of equity securities: Evidence from the ADR market. *The Journal of Finance* 58: 1583–1611.

Chakravarty, S., C. Chiyachantana, and C. Jiang. 2004. The choice of trading venue and relative price impact of institutional trading: ADRs versus the underlying securities in their local markets. Working paper, Purdue University, Singapore Management, and University of Memphis.

Christoffersen, S., C. Géczy, D. Musto, and A. Reed. 2005. Crossborder dividend taxation and the preferences of taxable and nontaxable investors: Evidence from Canada. *Journal of Financial Economics* 78: 121–144.

Cook, K. 2007. Stock price reaction to a reduction in the capital gains tax rate: Evidence from the Taxpayer Relief Act of 1997. Working paper, Texas A&M University.

Cowett, A. 2004. ADRs can hold corporate-governance surprises. *The National Law Journal*: 20–24.

Dai, Z., E. Maydew, D. Shackelford, and H. Zhang. 2008. Capital gains taxes and asset prices: Capitalization or lock-in? *The Journal of Finance* 63: 709–742.

De Jong, A., L. Rosenthal, and M. van Dijk. 2005. The limits of arbitrage: Evidence from dual-listed companies. Working paper, Erasmus University and Bentley College.

Dhaliwal, D., L. Krull, O. Li, and W. Moser. 2005. Dividend taxes and implied cost of equity capital. *Journal of Accounting Research* 43: 675–708.

———, ———, ———, and ———. 2007. Did the 2003 Tax Act reduce the cost of equity capital? *Journal of Accounting and Economics* 43: 121–150.

Domowitz, I., J. Glen, and A. Madhavan. 2001. Liquidity, volatility and equity trading costs across countries and over time. *International Finance* 4: 221–255.

Ellis, K., O. Li, and J. Robinson. 2006. Capital gains taxes and IPO underpricing. Working paper, University of California, University of Notre Dame, and The University of Texas at Austin.

Erickson, M. 1998. The effect of taxes on the structure of corporate acquisitions. *Journal of Accounting Research* 36: 279–298.

Errunza, V., and D. Miller. 2000. Market segmentation and the cost of capital in international equity markets. *Journal of Financial and Quantitative Analysis* 35: 577–600.

Eun, C., and S. Sabherwal. 2003. Cross-border listing and price discovery: Evidence from U.S. listed Canadian stocks. *The Journal of Finance* 58: 549–576.

Feldstein, M., J. Slemrod, and S. Yitzhaki. 1980. The effects of taxation on the selling of corporate stock and the realization of capital gains. *The Quarterly Journal of Economics* 94: 777–791.

Froot, K., and E. Dabora. 1999. How are stock prices affected by the location of trade? *Journal of Financial Economics* 53: 189–216.

Frost, C., and W. Kinney. 1996. Disclosure choices of foreign registrants in the United States. *Journal of Accounting Research* 34: 67–84.

Gagnon, L., and G. A. Karolyi. 2004. Multi-market trading and arbitrage. Working paper, Queen's University and The Ohio State University.

Grammig, J., M. Melvin, and C. Schlag. 2005. Internationally cross-listed stock prices during overlapping trading hours: Price discovery and exchange rate effects. *Journal of Empirical Finance* 12: 139–164.

Guenther, D., and M. Willenborg. 1999. Capital gains taxes and the cost of capital for small business: Evidence from the IPO market. *Journal of Financial Economics* 53: 385–408.

———, and R. Sansing. 2006. Fundamentals of shareholder tax capitalization. *Journal of Accounting and Economics* 42: 371–383.

Hail, L., and C. Leuz. 2009. Cost of capital effects and changes in growth expectations around U.S. cross-listings. *Journal of Financial Economics.*

Hubbard, J., and R. Michaely. 1997. Do investors ignore dividend taxation? A reexamination of the Citizens Utilities case. *Journal of Financial and Quantitative Analysis* 32: 117–135.

Karolyi, G. A. 1998. Why do companies list shares abroad? A survey of the evidence and its managerial implications. *Financial Markets, Institutions, and Instruments* 7: 1–60.

———. 2004. The role of ADRs in the development of emerging equity markets. *The Review of Economics and Statistics* 86: 670–690.

———. 2006. The world of cross-listings and cross-listings of the world: Challenging conventional wisdom. *Review of Finance* 10: 99–152.

Kim, M., A. Szakmary, and I. Mathur. 2000. Price transmission dynamics between ADRs and their underlying foreign securities. *Journal of Banking & Finance* 24: 1359–1382.

Klein, P. 1998. The capital gain lock-in effect with short sales constraints. *Journal of Banking & Finance* 22: 1533–1558.

———. 1999. The capital gain lock-in effect and equilibrium returns. *Journal of Public Economics* 71: 355–378.

Kothari, S., J. Sabino, and T. Zach. 2005. Implications of survival and data trimming for tests of market efficiency. *Journal of Accounting and Economics* 39: 129–161.

Koumkwa, S., and R. Susmel. 2005. Arbitrage and convergence: Evidence from Mexican ADRs. Working paper, University of Houston.

Lamont, O., and R. Thaler. 2003. Anomalies: The law of one price in financial markets. *The Journal of Economic Perspectives* 17: 191–202.

Landsman, W., and D. Shackelford. 1995. The lock-in effect of capital gains taxes: Evidence from the RJR Nabisco leveraged buyout. *National Tax Journal* 48: 245–259.

Lang, M., and D. Shackelford. 2000. Capitalization of capital gains taxes: Evidence from stock price reactions to the 1997 rate reduction. *Journal of Public Economics* 76: 69–85.

Leuz, C., and R. Verrecchia. 2000. The economic consequences of increased disclosure. *Journal of Accounting Research* 38: 91–124.

———. 2003. Discussion of ADRs, analysts and accuracy: Does cross-listing in the United States improve a firm's information environment and increase market value? *Journal of Accounting Research* 41: 347–362.

Maldonado, R., and A. Saunders. 1983. Foreign exchange futures and the law of one price. *Financial Management* 12: 19–23.

McDonald, R. 2001. Cross-border investing with tax arbitrage: The case of German dividend tax credits. *Review of Financial Studies* 14: 617–657.

Mehta, N. 2003. A better ADR mousetrap? Why the costs of trading are coming down. *Traders Magazine* 16: 68–70.

Miller, D., and M. Morey. 1996. The intraday pricing behavior of international dually listed securities. *Journal of International Financial Markets, Institutions and Money* 6: 79–89.

Miller, M., and M. Scholes. 1978. Dividends and taxes. *Journal of Financial Economics* 6: 333–364.

Patro, D. 2000. Return behavior and pricing of American depositary receipts. *Journal of International Financial Markets, Institutions and Money* 10: 43–67.

Pontiff, J. 1996. Costly arbitrage: Evidence from closed-end funds. *The Quarterly Journal of Economics* 111: 1135–1151.

Poterba, J., and L. Summers. 1985. *The Economic Effects of Dividend Taxation. Recent Advances in Corporate Finance*. Homewood, IL: Irwin Publishers.

Reese, W. 1998. Capital gains taxation and stock market activity: Evidence from IPOs. *The Journal of Finance* 53: 1799–1819.

Rosenthal, L., and C. Young. 1990. The seemingly anomalous price behavior of Royal Dutch/Shell and Unilever N.V./PLC. *Journal of Financial Economics* 26: 123–141.

Sarkissian, S., and M. Schill. 2008. Are there permanent valuation gains to overseas listing? *Review of Financial Studies* 22: 371–412.

Shackelford, D., and T. Shevlin. 2001. Empirical tax research in accounting. *Journal of Accounting and Economics* 31: 321–387.

———, and R. Verrecchia. 2002. Intertemporal tax discontinuities. *Journal of Accounting Research* 40: 205–222.

Wall Street Journal. 2004. Headwaters wins in arbitration. (April 20).

THE ACCOUNTING REVIEW
Vol. 84, No. 5
2009
pp. 1363–1394

American Accounting Association
DOI: 10.2308/accr.2009.84.5.1363

The Impact of Earnings on the Pricing of Credit Default Swaps

Jeffrey L. Callen
University of Toronto

Joshua Livnat
New York University

Dan Segal
Interdisciplinary Center Herzliya and University of Toronto

ABSTRACT: This study evaluates the impact of earnings on credit risk in the Credit Default Swap (CDS) market using levels, changes, and event study analyses. We find that earnings (cash flows, accruals) of reference firms are negatively and significantly correlated with the level of CDS premia, consistent with earnings (cash flows, accruals) conveying information about default risk. Based on the changes analysis, a 1 percent increase in ROA decreases CDS rates significantly by about 5 percent. We also find that (1) CDS premia are more highly correlated with below-median earnings than with above-median earnings and (2) CDS premia are more highly correlated with earnings of low-rated firms than with earnings of high-rated firms. Evidence indicates further that short-window earnings surprises are negatively and significantly correlated with CDS premia changes in the three-day window surrounding the preliminary earnings announcement, although the impact is concentrated in the shorter maturities.

Keywords: *credit default swaps; earnings; default risk; cash flows; accruals.*

JEL Classification: *M41, G13, G20, G32.*

Data Availability: *All data are publicly available.*

The authors acknowledge Lombard Risk for the CDS data provided to the Stern School Faculty through Professor Steve Figlewski, Charter Oak Investment Systems Inc. for the Point-In-Time Compustat quarterly data, and S&P's Compustat for the SEC Filing Dates database. We also acknowledge comments from workshop participants at the City University of Hong Kong, Hong Kong Polytechnic University, Interdisciplinary Center (Herzliya), Jiatong Xian University, Russian Academy of Sciences (St. Petersburg), Tsinghua University, and the University of Toronto. We are grateful to the anonymous reviewers of this journal and Steven Kachelmeier (senior editor) for comments and especially to Steve Huddart (editor) for both his insights and his willingness to help bring this project to fruition.

Editor's note: Accepted by Steven Kachelmeier, with thanks to Steven Huddart for serving as editor on previous versions.

Submitted: March 2007
Accepted: January 2009
Published Online: September 2009

I. INTRODUCTION

This study investigates the impact of earnings, accruals, and cash flows on firm credit risk as reflected in Credit Default Swaps (CDS). The CDS is essentially a "pure credit" default instrument, providing a far less noisy measure of credit risk in comparison to other debt instruments. Moreover, the credit derivative market, which is dominated by CDS contracts, is a *multi-trillion* dollar market (notional value) that has roughly doubled in size in the past five years. The existence of the CDS market and its burgeoning growth are prima facie evidence that other debt markets are unable to provide adequate solutions for the trading of credit risk. Thus, it is our contention that the CDS market is the best venue within which to investigate the extent to which earnings and its components are determinants of credit risk, and superior for this objective than the heterogeneous corporate bond and secondary loan markets.

We use three methodologies to examine the relation between CDS premia and earnings: a levels analysis, a changes (returns) analysis, and a short-window event study. A few contemporaneous studies (Benkert 2004; Batta 2006; Das et al. 2009) also incorporate some measure of earnings in a levels analysis of CDS rates. However, these studies do not examine the impact of accounting information on *changes* in CDS premia.[1] Changes and event study analyses provide more direct evidence concerning the credit risk information conveyed by earnings than does a levels analysis because the firm is its own control and heteroscedasticity concerns are mitigated. Furthermore, we find that levels of CDS premia, in contrast to changes, are uniformly nonstationary for all maturities and restructuring clauses. In addition, while Batta (2006) and Das et al. (2009) find that CDS rates and earnings levels are inversely related, Benkert (2004), using a much larger sample, finds a positive relation. This contradictory finding calls for a more comprehensive analysis using multiple methodologies.[2]

We also test for the potential asymmetry (nonlinearity) of CDS spreads to good news versus bad news, and to high- versus low-debt-rated firms. This issue is of interest to both accounting and finance researchers because prior research finds conflicting evidence of asymmetry in bond prices (e.g., Datta and Dhillon 1993; Easton et al. 2007). We also evaluate the impact of various CDS contract maturities on the earnings pricing relation. The relation between credit risk and maturity is potentially subtle (Ryan 2007, 96) and the evidence from the CDS market is relatively unexplored.[3]

Our analysis indicates that earnings, cash flows, and accruals have a statistically and economically significant impact on CDS premium. Using the levels (changes) analysis, we find that a 1 percent increase in ROA reduces CDS premia by close to 4.5 (2.8) basis points, or by about 9 (5) percent. Similarly, we find that the impact of 1 percent increase in cash flows scaled by total assets reduces CDS premia by 3.3 to 4.6 basis points, or 6 to 9 percent, whereas a similar change in accruals scaled by total assets reduces the premium by 2.7 to 4.2 basis points, or 5 to 8 percent.

We also find, as predicted, that the relations between CDS premia and each of earnings, cash flows, and accruals are significant for short to medium terms CDS, but are not significant for longer maturities. Further, the correlations decrease monotonically with maturity. In the changes analysis we find significant relations across most maturities.

---

[1] Batta (2006) reports a spread changes analysis in one of his tables (VII). However, the table does not report coefficient estimates, their statistical significance, or whether the independent variables are levels or changes.

[2] In contrast to other studies, this study uses the exact filing date of the financial statements as the relevant date for the CDS premia, thereby minimizing the possibility that CDS premia may incorporate other information unrelated to the information in the financial statements.

[3] Das et al. (2008) control for maturity but, again, do not account for potential nonlinearities.

Furthermore, both the levels and changes analyses provide evidence consistent with nonlinear asymmetric payoff functions for debt holders. Specifically, CDS premia are more highly correlated with below-median earnings than with above-median earnings, and CDS premia are more highly correlated with the earnings of low-rated firms—BBB and below—than the earnings of high-rated firms.

The analysis of CDS premia across the three-day window surrounding preliminary earnings announcements indicates that CDS premia are inversely and significantly related to earnings surprises, especially for shorter maturities. CDS premia are also significantly related to the non-earnings information contained in cumulative equity returns over the same window. Again, short-window earnings changes convey CDS value-relevant information, primarily for firms with lower credit ratings.

Section II briefly describes the CDS market and indicates the reasons why the CDS market is superior to the corporate bond and the secondary loan markets for determining the relation between earnings and credit risk. Section III surveys the relevant CDS literature. Section IV develops the hypotheses to be tested. Section V describes the sample. Section VI presents the empirical analysis. Section VII briefly relates the findings of this study to the current global credit crisis and concludes.

II. THE CDS MARKET

Basics

Financial innovation led to the creation of credit derivatives to manage credit risk. The extraordinary growth in credit derivatives can be attributed to three main market participants. The largest group is comprised of large commercial banks, which are net buyers of credit derivative protection. Credit derivatives enable banks to transfer and diversify the credit risk of their loan portfolios but, in contrast to securitizations, without removing the loans from bank balance sheets and without directly involving borrowers. Insurance companies, the second largest group, are net sellers of credit derivatives, as they attempt to enhance investment yields by diversifying their exposure to risks that are uncorrelated with the existing insurance business. Global hedge funds have also become large players in the credit derivative market, operating both as protection buyers and sellers in an effort to carry out complex arbitrage strategies across various financial markets. Other miscellaneous participants include financial guarantors, which are net protection sellers, and pension funds.

The most common credit derivative is the single-name Credit Default Swap. The CDS market has grown rapidly, in no small part due to the standardization of CDS contracts by the International Swaps and Derivatives Association. A CDS is an over-the-counter contract between two parties that provides protection against credit risk, where the protection buyer pays a fixed premium, the spread, to the protection seller for a period of time. If a certain pre-specified credit event occurs to a specific company, called the reference entity, then the protection seller pays compensation to the protection buyer. The reference entity is not a party to the contract, and it is not necessary for the buyer or seller to obtain the reference entity's consent to enter into a CDS. Credit events in CDS contracts often include failure to pay, bankruptcy, and restructuring of the reference entity. If the protection buyer does not hold the reference bond, then compensation is in the form of a lump sum payment equal to the difference between the value of the reference entity's bond and its face value. Alternatively, if the protection buyer holds the bond, then the protection buyer either receives the cash difference or delivers the reference entity's bond to the protection seller for its face value. If no credit event occurs during the term of the swap, then the protection buyer continues to pay the swap premium until maturity. Maturities typically range from one to ten years, with the five-year maturity being the most common.

The premium paid by the protection buyer to the seller is quoted in basis points per annum of the contract's notional value—typically $10 million—and is usually paid quarterly. For example, suppose that the CDS spread for a five-year $10 million contract on British Airways is 125 basis points and the credit event is defined to be default of British Airway's bonds. This means that the protection buyer pays $31,250 (1.25/4 percent of $10M) per quarter to the protection seller for five years or until British Air defaults on its bonds. In the event that British Air defaults on its bonds, the protection buyer, in turn, has the right to sell British Air's bonds to the protection seller for their face value of $10 million.

A number of institutional differences between corporate bonds (and secondary loans) and the CDS market have contributed to the phenomenal growth in CDS trading. A CDS does not require initial funding, which allows for leveraged positions. A CDS transaction can be entered into even if a cash bond of the reference entity at a particular maturity is unavailable. Finally, by entering a CDS contract as a protection seller, an investor can easily create a short position in the reference entity's credit.

CDS Instruments versus Corporate Bonds and Secondary Loans

The credit-risk profile of a CDS is similar to that of a corporate bond of the reference entity. In fact, in the absence of (1) arbitrage opportunities, (2) contractual features such as embedded options, covenants, and guarantees, and (3) market frictions, the CDS premia and the corporate bond yield spread are identical for a floating rate corporate bond (Duffie 1999).

Nevertheless, CDS premia offer many advantages over corporate bond yield spreads and secondary loan spreads for analyzing the determinants of credit risk. First, bond spreads include factors unrelated to credit risk, such as systematic risk unrelated to default (Elton et al. 2001) and illiquidity (Longstaff et al. 2005). In fact, Huang and Huang (2002) conclude that less than 25 percent of the credit spread in corporate bonds is attributable to credit risk. Second, interest risk *drives* fixed-rate corporate bond yields and secondary market loan rates quite independently of credit risk. A positive relation between earnings and bond prices could well be due to changes in interest rates, if, say, risk-free interest rate decreases cause firm earnings to increase over time, rather than to credit risk. Third, in contrast to CDS instruments, corporate bonds and secondary market loans are replete with embedded options, guarantees, and covenants. Heterogeneity in these features potentially distorts the relationship between earnings and credit risk in cross-sectional studies. Even more problematic is that they may generate a spurious relation between earnings and credit risk. For example, the positive relation between earnings and corporate bond prices could be driven by earnings-based covenants rather than by credit risk, per se. With lower earnings, earnings-based covenants are more likely to be binding, increasing the probability of technical bankruptcy and concomitant expected transactions (renegotiation) costs and, thereby, leading to reduced bond prices. In contrast, except in rare cases, technical default is not a credit event in CDS contracts. Fourth, the available empirical evidence indicates that credit risk price discovery takes place first in the CDS market and only later in the bond market (Blanco et al. 2005; Zhu 2006; Daniels and Jensen 2005). The bond market's lagged reaction potentially distorts empirical studies relating earnings to bond prices. Fifth, unlike corporate bond yield spreads, no benchmark risk-free rate need be specified for CDS premia, minimizing potential misspecification of the appropriate risk-free rate proxy.[4] Finally, CDS rates are closely related to the par value of the reference bond, whereas corporate bond values (including their taxability characteristics) are affected by coupons. Heterogeneity in coupon rates potentially distorts the relationship between earnings and credit risk in cross-sectional studies.

III. CDS LITERATURE SURVEY

The finance literature proposes three model types to explain the pricing of credit derivates: structural models, reduced form models, and hybrid models. Based upon the seminal study by Merton (1974), structural models imply that the main determinants of the likelihood and severity of default are financial leverage, the volatility of the firm's assets, and the risk-free rate of interest. Because structural models assume that (1) investors observe the reference entity's assets directly and (2) the reference entity's asset structure behaves according to a known stochastic process, neither earnings nor other accounting information play a role in the pricing of CDS in these models. Thus, from the perspective of these models, the null hypothesis that earnings play no part in the pricing of credit risk (as reflected in CDS premia) is quite credible.

Reduced-form models (Das 1995; Das and Sundaram 2000; Hull and White 2000a, 2000b) do a better job at directly pricing credit derivatives. Nevertheless, reduced form models leave no room for earnings as determinant of credit risk because they exogenously postulate the dynamics of default probabilities. Thus, reduced form models tend to be "black boxes" that are silent about the underlying determinants of the price of credit risky instruments.

Far better economic insights are provided by the hybrid model of Duffie and Lando (2001), which assumes that the rate of default is a function of the firm's asset value and capital structure. Unlike standard structural models, investors cannot observe the asset structure of the firm directly. Instead, investors receive periodic accounting reports that provide imperfect information about the true financial condition of the firm. Accounting information in the Duffie and Lando (2001) model is quite abstract. Except for the notion of noisy accounting information about the firm's asset structure, also referred to by the authors as accounting information transparency, no specific accounting variables are specified in the model. However, as Duffie and Lando (2001) point out in the generalization section of their article, their model can be readily extended to accommodate other types of accounting information, including accounting ratios and peer performance measures. Therefore, besides the standard structural model variables of conventional structural models, the hybrid model of Duffie and Lando (2001) also rationalizes accounting information transparency, earnings, cash flows, and accruals, as well as accounting information generally, as determinants of CDS pricing.

Rather then trying to directly price credit derivatives using a structural or hybrid model, we follow the regression approach of Collin-Dufresne et al. (2001, hereafter CGM). They "borrow" from the structural approach to identify the theoretical determinants of credit risk. These determinants are then used as independent variables to explain corporate credit spreads, rather than inputs to a particular structural model. Although CGM themselves estimate bond yield spreads, a similar approach using CDS premia is taken by Aunon-Nerin et al. (2002), Benkert (2004), Abid and Naifar (2006), Ericsson et al. (2006, 2008), Batta (2006), and Das et al. (2009). Three of these empirical studies, two of which are essentially contemporaneous with ours, investigate the pricing implications of earnings on CDS prices using levels analyses. Benkert (2004) regresses *daily* five-year CDS premia on earnings to sales and earnings to interest, as well as on other determinants of CDS spreads such as leverage and volatility. Benkert (2004) obtains the counterintuitive result that, controlling for credit ratings, the earnings variables are significantly and *positively* related to CDS premia.[5] Batta (2006) regresses levels of five-year CDS premia on interest coverage ratios (i.e., earnings plus interest expense divided by interest expense) and other control variables. In contrast to Benkert (2004), Batta (2006) finds that interest coverage is significant and negatively related to CDS premia, primarily when the coverage ratio is non-positive. More recently, Das et al. (2009) implement a (regression) horse race between a market-based model of CDS rates and an accounting-based model of CDS rates inclusive of an earnings variable. They find that a levels model (using pooled data) based solely on accounting-based metrics—primarily ratios and growth rates—does as well in terms of R^2 as a pure market-based model and, in particular, they find that earnings is inversely related to CDS spreads.

---

[4] See Houweling and Vorst (2005) on this issue. Another disadvantage of bond yield spreads includes the tax differentials in bond pricing. Elton et al. (2001) document a tax premium of 29 to 73 percent of the corporate bond spread, depending on the rating. See Longstaff et al. (2005) for other potential disadvantages of bond spreads as measures of credit risk relative to CDS premia.

IV. HYPOTHESES DEVELOPMENT

The hybrid model of Duffie and Lando (2001) partially motivates the hypotheses of this study. The price of a CDS instrument in their model is a function of (1) the factors that explain CDS pricing in standard structural models and (2) the imperfect information available to the CDS market about the firm's asset dynamics from periodic corporate financial reports. Although accounting information in the Duffie-Lando model is completely generic, we focus on earnings for two reasons. First, earnings are (arguably) the most important information variable used by investors to evaluate firm performance and, hence, future firm wealth, a major determinant of credit risk. Second, earnings can be used by investors to estimate the reference entity's true asset dynamics and, hence, its credit risk. More specifically, increased profitability of the reference entity, as measured by current accounting earnings, should reduce its credit risk since, with increased profitability, the reference entity is wealthier and less likely to default. Accounting studies have shown that current earnings are a good predictor of future earnings (Finger 1994; Nissim and Penman 2001), future cash flows (Dechow et al. 1998; Barth et al. 2001), and firm equity performance (Dechow 1994). In other words, an increase in earnings portends an increase in operating and equity performance and, hence, a reduced probability of bankruptcy. Also, earnings comprise a significant portion of the short-term change in firm assets (via clean surplus) and, therefore, provide information to investors about the firm's asset dynamics.

---

[5] We find a significant and negative relation between earnings and CDS premia for our sample even after controlling for ratings in the same fashion as Benkert (2004). We further investigate the disparity between Benkert's (2004) results and ours by replicating his study on our sample. His sample period is from 1999 through May 2002. When we restrict our sample to Benkert's sample period, we find that the coefficients on earnings to sales and earnings to interest are not statistically different from zero. One potential explanation for Benkert's finding is that he uses firm fixed effects in the regression. Since credit ratings are sticky, there is a very high correlation between the firm fixed effects and credit ratings. The problem is exacerbated by using daily CDS prices as does Benkert. Indeed, when we estimate Benkert's regression specification using a random-effects model, we find that the coefficient on earnings to sale is negative and significant. One caveat is that our sample for Benkert's period is far smaller than his; we have about 3,000 observations, whereas Benkert has more than 26,000. Thus, it is possible that the difference between our results and Benkert's are driven by sample size. When we estimate Benkert's regression using our entire sample (1999–2005, 285,275 observations) we find that the coefficient on earnings to sales (earnings to interest) is negative (positive) and significant. These results hold for the firm random-effects model as well.

Furthermore, the shorter the maturity of the CDS contract at initiation, the more likely is an increase (decrease) in profitability to provide positive (negative) information about the reference entity's ability to repay bondholders within the maturity period. Conversely, longer term (initial) maturities are less likely to be affected by current earnings, because earnings are less informative about long-term performance than short-term performance. Potential nonlinearities in the credit risk maturity relation, however, can offset these considerations. For example, credit risk can be positively (negatively) related to maturity for less (more) risky borrowers (Ryan 2007). Ultimately, the relation between CDS rates, earnings, and maturity is an empirical question.

These considerations yield the following set of hypotheses, stated in the alternative:

H1: The quarter-end CDS premium is inversely related to the reference entities' quarterly earnings; the relation between the variables is a decreasing function of CDS maturity.

H2: The percent change in the CDS premium over the quarter is inversely related to the *change* in the reference entities' quarterly earnings; the relation between the variables is a decreasing function of CDS maturity.

H3: The percent change in the CDS premium in the short window centered on the preliminary earnings announcement is inversely related to the reference entities' earnings surprise; the relation between the variables is a decreasing function of CDS maturity.

There appears to be a consensus among accounting scholars that both accruals and cash flows are value-relevant in equity markets. In particular, a large number of studies find that both accruals and cash flows are contemporaneously related to equity returns and current accruals and current cash flows can be used to predict future equity prices (e.g., Sloan, 1996; Cheng and Thomas 2006; Livnat and Lopez-Espinosa 2008). This literature suggests that, similar to earnings, accruals and cash flows convey information about the reference entity's future wealth and about its true asset dynamics and, hence, its credit risk. However, unlike cash flows, accruals are subject to manipulation, suggesting that accruals are less predictive of firm wealth than cash flows. Moreover, higher accruals may reflect higher credit risk because there may be less cash to satisfy the claims of creditors for a given level of earnings. Hence, accruals are either negatively related to credit risk, but less so than cash flows, or accruals are positively related to credit risk.

H4: The (change in the) CDS premium is inversely related to the (change in the) reference entities' cash flows. The (change in the) cash flows have (has) a greater negative association with credit risk than accruals.

As we have argued, earnings can convey credit-risk-relevant information because earnings are informative about the current wealth of the firm, the future operating and equity performance of the firm, and the dynamics of the asset structure. Viewing the equity of the firm as an option on the firm's assets suggests that debt holders will have a nonlinear payoff function: namely, they will react more to information that presages potential bankruptcy than information that presages additional profits. One possible means of testing for this nonlinearity is to see whether there is an asymmetry in the response of CDS premia to losses versus profits. Extant evidence of such nonlinearities in the corporate bond market is mixed. Datta and Dhillon (1993) find that corporate bond yields do *not* react more to (unexpected) losses than to (unexpected) profits, whereas Easton et al. (2007) and DeFond and Zhang (2008) find such an asymmetry. Since CDS reference firms are typically quite large and rarely incur losses, we examine whether CDS spreads react more to earnings of firms with below the three-digit industry median profitability than to earnings of firms with above industry-median profitability.

Another means of testing for nonlinearity in CDS premia is through credit ratings. Specifically, earnings should be more highly associated with CDS premia for those firms for which the likelihood of bankruptcy is greater, as measured by credit ratings. To proxy for the likelihood of bankruptcy, we divide firms into two groups based on credit ratings.[6] Firms with credit ratings BBB and below (splitting the sample into two equal groups) are assumed to be more likely to go bankrupt than firms with higher ratings. This classification is also consistent with the CDS literature that maintains that CDS prices are nonlinear in the reference entities' credit ratings (e.g., Aunon-Nerin et al. 2002). These considerations lead to our last hypothesis.

H5: The relation between (changes in) each of earnings, accruals, and cash flows and (percentage changes in) CDS premia is greater in absolute value for firms with low credit ratings (below median earnings) than for firms with high credit ratings (above median earnings)

V. DATA

CDS data for 2002 through 2005 are obtained from Lombard Risk. For a given date and reference firm, this database reports a composite at-market CDS rate for the initial maturity. This composite rate is derived from the mid-market quotes contributed by up to 25 investment banks and default-swap brokers. All quotes are for the initial maturity only. Besides quotes, the database includes, for each date, the referenced firm identifier (primarily ticker symbol), the seniority (senior or subordinated), the currency of the underlying debt, the initial maturity of the CDS contract (1, 3, 5, 7, or 10 years), the standard deviation of the mid-market quotes, and the restructuring clause applied in the contract. Also reported is the average recovery rate used by the quote providers. The frequency of observations has increased over time. The database contains biweekly information from January 2002 to June 2002, weekly data from July 2002 to May 2003, and daily quotes starting May 15, 2003. The initial sample comprises 2,127,526 CDS contracts for 798 firms.

Using the Preliminary Compustat database available to us through Charter Oak (and now available through WRDS), we identify for each firm on Compustat the quarterly preliminary earnings release date. In addition, we obtain SEC filing dates for sample firms from S&P's SEC Filing Dates Database (also available through WRDS). We use ticker symbols to merge the Compustat and SEC filing dates data with the CDS data set. Requiring CDS data around the preliminary earnings release date or around the financial statement filing date results in a sample of 598 (4,890) firms (firm-quarters) and 46,602 CDS contracts. To enhance the homogeneity of the sample, we eliminate contracts with non-seniority clauses (1,127 contracts), contracts with Modified-Modified Restructuring clauses (1,024 contracts), and contracts not denominated in U.S. dollars (7,382 contracts). These restrictions reduce the sample size to 37,069 contracts for 577 firms (4,837 firm-quarters). We further require that each observation has sufficient data to compute either the change in CDS premium around the preliminary earnings release date (for the event study) or the change in CDS premium in the most recent quarter or the CDS premium one day after the SEC filing date, resulting in a sample of 21,839 contracts for 536 firms (3,089 firm-quarters). Merging this sample with Compustat and CRSP and requiring non-missing earnings, market value of equity greater than $100 million, and positive book value of equity, results in a final sample of 20,328 contracts for 508 firms (2,800 firm-quarters). Table 1 describes the data and sample filters. Except for Benkert (2004), who uses daily data, this sample is significantly larger than those of the prior CDS studies referenced above.

---

[6] To proxy closeness to default, Berndt et al. (2006) use Merton's (1974) "distance to default" measure defined roughly as the number of standard deviations of asset growth by which a firm's market value of assets exceeds a liability measure. This measure is highly correlated with stock return volatility, an important variable in our analysis. We use high/low credit ratings instead as a proxy of closeness to default.

Table 2 presents descriptive statistics for our sample observations. The firms covered in the database are large; the mean (median) firm has a market value of $13.09 (6.70) billion. Thus, the results reported in this study are unlikely to generalize to smaller firms. The percentage change of the CDS premium in the three-day window centered on the preliminary earnings announcement date has a mean of 0.7 percent and a median of 0.0 percent. The percentage change in the CDS premium from one day after the SEC filing date of the previous quarter through one day after the SEC filing date of the current quarter, *CDS_QTR*, has a mean (median) of 0.0 (−7) percent, suggesting that the median CDS premium tends to decrease with the additional information learned throughout the quarter. The log of the CDS premium on the first day after the SEC filing date, *CDS_PRM*, has an average (median) value of 4.08 (3.95), representing a premium of 59 (52) basis points.

The mean (median) quarterly return on assets (*ROA*), defined as income before extraordinary items and discontinued operations scaled by total assets at the end of the quarter, is 1.2 (1.1) percent, which is reasonable considering the large firm bias in the sample. The

TABLE 1
Data Filters

Sample Structure	Observations
Number of observations in the CDS daily dataset with non-missing maturity and premium (2002–2005)	2,127,526 daily observations of 798 firms
Interaction of CDS daily dataset and the preliminary Compustat data (note that each firm quarter usually contains a number of CDS contracts that differ in their maturity, restructuring clauses, etc.)	46,602 CDS contracts, 4,890 firm-quarters of 598 firms
Excluding subordinated CDS contracts	45,475 CDS contracts, 4,854 firm-quarters of 582 firms
Excluding CDS contracts with XMMR restructuring clause	44,451 CDS contracts, 4,847 firm-quarters of 580 firms
Excluding CDS contracts denominated in currency other than U.S.$	37,069 CDS contracts, 4,837 firm-quarters of 577 firms
Requiring that each observation has either CDS premium a day after the filing date, or CDS premium around the preliminary earnings release date	21,839 CDS contracts, 3,089 firm-quarters of 536 firms
Merging the data with the Compustat database for 2002–2005	21,022 CDS contracts, 2,905 firm-quarters of 526 firms
Eliminating firms with MV below $100 million and negative book value of equity	20,328 CDS contracts, 2,800 firm-quarters of 508 firms

TABLE 2
Descriptive Statistics

Variable	n	Mean	SD	Q1	Median	Q3
MV	20328	13,088	22,900	3,098	6,696	14,083
DEL_CDS_PRL	14044	0.007	0.144	−0.030	0.000	0.022
DEL_CDS_QTR	10992	0.000	0.448	−0.239	−0.070	0.117
CDS_PRM	18222	4.080	1.046	3.360	3.951	4.723
ROA	20328	0.012	0.023	0.004	0.011	0.021
DEL_ROA	20310	0.002	0.029	−0.002	0.001	0.005
LEV	20328	0.446	0.203	0.295	0.425	0.576
DEL_LEV	20328	−0.006	0.039	−0.026	−0.007	0.013
ACC	19828	−0.012	0.030	−0.022	−0.011	−0.001
OCF	19839	0.024	0.027	0.010	0.022	0.037
RATE	14936	10.395	3.006	9.000	11.000	12.000
EAR_SUR	19601	0.001	0.009	0.000	0.000	0.002
SD_RET	20318	0.018	0.009	0.013	0.016	0.022
SPOT	20328	2.285	0.910	1.270	2.120	3.320

This table provides descriptive statistics of the main variables used in the paper.

Variable Definitions:

MV = market value (in $MM) of equity at the end of fiscal quarter;
DEL_CDS_PRL = change in CDS premium in the three-day window centered on the preliminary earnings release date. It is computed as the CDS premium on the last day of the window divided by the CDS premium on the first day of the window minus 1;
DEL_CDS_QTR = change in the CDS premium during the quarter. It is computed as the premium on the first day after the SEC filing date of the current period divided by the premium on the first day after the SEC filing date of the previous quarter minus 1;
CDS_PRM = log of the CDS premium (in basis points) on the first day after the SEC filing date;
ROA = the return on assets, computed as quarterly net income before extraordinary items divided by total assets;
DEL_ROA = change in *ROA* and it is computed as *ROA* for current quarter minus *ROA* in the same quarter of the previous year;
LEV = leverage, computed as long-term debt scaled by the value of assets (market value of equity + book value total liabilities);
DEL_LEV = change in leverage;
ACC (*OCF*) = total quarterly accruals (operating cash flows) scaled by total assets in the previous quarter. Accruals are net income before extraordinary items minus net operating cash flows;
RATE = S&P short-term credit rating;
EAR_SUR = earnings surprise; it is calculated as actual I/B/E/S earnings minus the median analysts' forecast, scaled by price at quarter-end;
SD_RET = standard deviation of daily returns during the firm's current fiscal quarter; and
SPOT = one-year T-Bill rate.

change in *ROA* from the previous quarter has a mean (median) of 0.2 (0.1) percent, indicating that most firms exhibit positive earnings growth, which is consistent with the negative percentage change in the median CDS premium over the quarter. Leverage (*LEV*), estimated as total debt divided by market value of assets (computed as market value of equity plus total liabilities) at quarter end, has a mean (median) of 0.45 (0.43). The change in *LEV* over the quarter has a mean of −0.006, indicating a slight decrease in leverage. The quarterly net operating cash flow (accrual) scaled by total assets at the previous quarter-end has a mean (median) of 2.4 (−1.2) percent, which is in line with prior studies. The S&P senior debt rating for the sample observations (*RATE*) shows considerable variation, with ratings of AAA (code 2) through B− (Code 18) with a median rating of BBB (code 11). The earnings surprise (*EAR_SUR*) is computed as actual I/B/E/S EPS minus the most recent median forecast, scaled by share price. Table 2 shows that the mean earnings surprise for our sample firms is 0.001. Other variables that are important in determining CDS premia (the volatility of returns and the risk-free interest rate) show considerable variation during our sample period. Finally, untabulated results show that 17 percent, 22 percent, 23 percent, 20 percent, and 18 percent of the sample CDS contracts have initial maturities of 1, 3, 5, 7, and 10 years, respectively.

VI. RESULTS

Control Variables

The tests in this study are based primarily upon cross-sectional analysis. Thus, it is crucial to control for factors other than earnings that are known to affect swap spreads. The main firm-level determinants of the likelihood and severity of default in structural and hybrid models are financial leverage, the volatility of the firm's assets, the riskless rate of interest, and the (initial) maturity of the contact. More specifically, the higher the leverage and the more volatile the assets of the reference entity, the higher is the probability of a credit event and so the higher is the swap premium. In contrast, the higher the riskless rate of interest, the lower is the swap premium. Intuitively, a higher rate of interest increases the firm's wealth and makes bankruptcy less likely. The swap premium increases with the initial maturity of the contract because longer contract maturities increase the probability that a credit event will occur. We also control for credit ratings and the restructuring provisions of the CDS contract.[7] Credit ratings provide a noisy measure of credit risk and are employed in all the empirical CDS studies referenced above. We expect that the lower the credit rating of the reference entity, the higher is the swap premium.

Our tests involve firm random effects with year fixed effects panel-data regressions, for which the dependent variables are either levels or (percentage) changes in CDS premia and the independent variables are (changes in) earnings, and other model-driven and control variables specified above that are potentially important in explaining CDS premia.[8] Because the data include various CDS contracts for the same firm, significance of the coefficients is based on firm-clustered (Rogers) standard errors, following Petersen (2009). To reduce potential multicollinearity induced by interaction terms, we demean all continuous independent variables (Aiken and West 1991).

Levels Analysis

Table 3 presents the levels analysis (H1). The *BASE* column of Table 3, Panel A provides the baseline regression derived from structural models of CDS pricing. Except for the restructuring controls, *CUMR* and *EXR* (defined below), this regression is common to virtually all of the CDS studies cited. In the *BASE* regression, the log of the level of CDS premia one day after the SEC filing date is regressed on leverage (*LEV*), the risk-free rate of interest (*SPOT*), the volatility of returns (*SD_RET*), and the S&P credit rating (*RATE*). The indicator variables *CUMR* and *EXR* denote full restructuring and no restructuring, respectively. *Inter alia*, the intercept captures modified restructuring. The indicator variables *D_3* through *D_10* denote CDS maturities of three through ten years. Thus, the intercept captures CDS of one-year maturity as well.

---

[7] In contradistinction to the three earnings-related CDS levels studies surveyed above, we control for restructuring clauses since recent empirical and theoretical work by Packer and Zhu (2005) and Berndt et al. (2006) indicate that restructuring clauses are important determinants of CDS pricing.

[8] Breusch-Pagan Lagrange Multiplier tests reject firm fixed effects in favor of firm random effects in all regressions.

TABLE 3
Levels Regression

Panel A: Regressions of CDS Premium on Earnings, and Cash Flow and Accruals[a]

	BASE	*ROA*	*CF_ACC*
Intercept	2.863***	6.261***	6.299***
	(0.000)	(0.000)	(0.000)
ROA		−3.242***	
		(0.005)	
OCF			−3.605***
			(0.002)
ACC			−2.921**
			(0.017)
LEV	2.343***	1.155***	1.136***
	(0.000)	(0.000)	(0.000)
SPOT	−0.134***	−0.118***	−0.116***
	(0.000)	(0.000)	(0.000)
SD_RET	16.067***	15.768***	15.708***
	(0.000)	(0.000)	(0.000)
RATE	0.046**	0.037**	0.037**
	(0.025)	(0.050)	(0.049)
CUMR	0.008	0.005	0.005
	(0.510)	(0.681)	(0.689)
EXR	−0.079***	−0.076***	−0.076***
	(0.000)	(0.000)	(0.000)
D_3	0.261***	0.261***	0.261***
	(0.000)	(0.000)	(0.000)
D_5	0.466***	0.467***	0.467***
	(0.000)	(0.000)	(0.000)
D_7	0.564***	0.565***	0.565***
	(0.000)	(0.000)	(0.000)
D_10	0.682***	0.682***	0.682***
	(0.000)	(0.000)	(0.000)
SIZE		−0.328***	−0.331***
		(0.000)	(0.000)
Adj. R^2	0.471	0.530	0.530
Chi-square	2,409***	2,586***	2,607***
n		12,242 contracts; 423 firms	

Panel B: Levels Regressions with Maturity Interaction[b]

	ROA	*CF_ACC*
Intercept	6.234***	6.263***
	(0.000)	(0.000)
ROA	−5.265***	
	(0.000)	
ROA_3	−0.387	
	(0.604)	
ROA_5	1.688*	
	(0.088)	

(*continued on next page*)

TABLE 3 (continued)

	ROA	*CF_ACC*
ROA_7	3.785*** (0.001)	
ROA_10	6.210*** (0.000)	
OCF		−5.361*** (0.000)
OCF_3		−0.633 (0.383)
OCF_5		1.344 (0.168)
OCF_7		3.423*** (0.002)
OCF_10		6.023*** (0.000)
ACC		−5.685*** (0.001)
ACC_3		−0.047 (0.960)
ACC_5		2.675** (0.034)
ACC_7		5.037*** (0.000)
ACC_10		7.415*** (0.000)
LEV	1.153*** (0.000)	1.135*** (0.000)
SPOT	−0.120*** (0.000)	−0.118*** (0.000)
SD_RET	15.811*** (0.000)	15.752*** (0.000)
RATE	0.037** (0.049)	0.037** (0.048)
CUMR	0.005 (0.683)	0.005 (0.692)
EXR	−0.075*** (0.000)	−0.075*** (0.000)
D_3	0.263*** (0.000)	0.264*** (0.000)
D_5	0.469*** (0.000)	0.469*** (0.000)
D_7	0.566*** (0.000)	0.567*** (0.000)
D_10	0.685*** (0.000)	0.685*** (0.000)
SIZE	−0.330*** (0.000)	−0.333*** (0.000)
Adj. R^2	0.531	0.531
Chi-square	3,230***	3,289***
n	12,242 contracts; 423 firms	

(*continued on next page*)

TABLE 3 (continued)

Panel C: Tests for Earnings, Accruals, and Cash Flows across Maturities[c]

	ROA	Accruals	Cash Flows
1-Year CDS	−5.265*** (0.000)	−5.685*** (0.000)	−5.361*** (0.000)
3-Year CDS	−5.651*** (0.000)	−5.733*** (0.000)	−5.994*** (0.000)
5-Year CDS	−3.577*** (0.000)	−3.010** (0.012)	−4.017*** (0.000)
7-Year CDS	−1.480 (0.212)	−0.648 (0.600)	−1.938 (0.105)
10-Year CDS	0.945 (0.424)	1.730 (0.163)	0.663 (0.580)

---

*, **, *** Indicates statistical significance at the 10 percent, 5 percent, and 1 percent levels, respectively.

[a] Panel A shows the panel data regressions results of CDS premia on earnings, and on cash flows and accruals. The dependent variable is the (log of) CDS premium one day after the SEC filing date. The regressions are estimated using firm random effects with year fixed effects (suppressed). We use robust standard errors clustered by firm. The two-tailed p-values of the coefficients are in parentheses.

[b] Panel B extends the results of the regressions in Panel A above by interacting earnings (*ROA*), and accruals and cash flows with the maturity indicator variables. The *ROA* coefficients with suffix *i* (*i* = 3,5,7,10) represent the interaction of *ROA* with *D_i*, an indicator variable equal to 1 if the CDS contract maturity is *i* years and 0 otherwise (*i* = 3,5,7,10). Thus, *ROA_i* represents the incremental effect of *ROA* on the premium of the CDS contract with maturity *i* over the impact of *ROA* on the premium of the CDS contract with one year to maturity. The *CF_ACC* column shows the results when we replace *ROA* with accruals and cash flows, respectively. To avoid potential multicollinearity and to be able to interpret the interactions correctly, *ROA*, *ACC* and *OCF* are demeaned prior to creating the interaction variables. The regressions are estimated using firm random effects with year fixed effects (suppressed). We use robust standard errors clustered by firm.

[c] Panel C shows the overall coefficients for *ROA*, *ACC*, and *OCF* for the different maturities. For example, the *ROA* coefficient on the three-year CDS is equal to the *ROA* coefficient plus the *ROA_3* coefficient from Panel B. The two-tailed p-values of the coefficients are in parentheses.

Variable Definitions:

CUMR (EXR) = an indicator variable with 1 if the CDS contract has full restructuring clause (exclude restructuring) and 0 otherwise;

D_i = an indicator variable with 1 if the maturity of the contract equals *i*, and 0 otherwise, where *i* = 3,5,7,10;

SIZE = the log of market value; and

n = number of observations (CDS contracts and firms).

All other variables are defined in the notes of Table 2.

The highly significant *BASE* levels regression yields results that are consistent with the underlying theory of structural models. The coefficients for leverage and the volatility of returns are positive and highly significant ($p < 0.001$, two-tailed). The coefficient for the risk-free rate of interest is negative and highly significant ($p < 0.001$, two-tailed). The credit rating—the higher is the *RATE* variable, the lower the credit rating—is positive and significant ($p = 0.025$, two-tailed). As expected, CDS contracts without restructuring clauses (*EXR*) have significantly lower CDS premia than CDS contracts with restructuring clauses ($p < 0.001$, two-tailed). The insignificance of the *CUMR* coefficient indicates that there is no significant difference between full and modified restructuring clauses. Further, CDS premia increase significantly and monotonically with (initial) contract maturity ($p < 0.001$, two-tailed), consistent with the option-pricing theory underlying CDS contracts.

The *ROA* column includes earnings normalized by total assets (*ROA*) and firm size (*SIZE*) in addition to the structural model variables and bond ratings. *SIZE* is a proxy for accounting information certainty/transparency. The larger the firm, the less uncertain/more transparent is the accounting information. The inclusion of these latter variables is consistent with the hybrid model of Duffie and Lando (2001).[9] As expected, both the earnings and size coefficients are negative and highly significant ($p < 0.010$, two-tailed). The remaining coefficients are qualitatively similar to those of the *BASE* regression.

The *CF_ACC* column replaces *ROA* after decomposing earnings into its accrual/cash flow components. These components are scaled by total assets. Consistent with H4, both the accruals and cash flow components are negative and highly significant ($p < 0.020$ and $p < 0.010$, respectively, two-tailed). Although the accruals coefficient is smaller in absolute value than the cash flow coefficient, suggesting that accrual earnings are less credible to investors than cash flows, the difference is not significant based on clustered standard errors, but is significantly different using conventional robust standard errors (p-value = 0.002). The remaining coefficients are qualitatively similar to those from the *BASE* and *ROA* regressions.

The impact of earnings on CDS premia is likely to depend upon the (initial) maturity of the CDS contract—the shorter the maturity, the greater should be the impact of earnings on CDS premia. Panel B of Table 3 replicates the *ROA* and *CF_ACC* levels regressions of Panel A after interacting earnings, accruals, and cash flows with the maturity indicator variables, respectively. *ROA_i* represents the *incremental* effect of earnings on the CDS premium of maturity contract *i*, relative to the one-year maturity (*ROA*) contract, $i = 3,5,7,10$. (Similar notation is used for cash flows and accruals.) The ROA column regression in Panel B shows that increased earnings reduces CDS premia significantly for one-year maturities ($p < 0.001$, two-tailed). The *marginal* interaction coefficients are positive and significant (with the exception of the three-year maturity) and increase monotonically with maturity, as predicted, indicating weaker association as maturity increases. Similar monotonic relationships hold for both cash flows and accruals in the *CF_ACC* column. All other coefficients are qualitatively similar to those of the Panel A regressions.

Panel C of Table 3 shows the total (instead of marginal as in Panel B) maturity effects, computed by adding the marginal coefficient to the base one-year coefficient. The *ROA* column of Panel C indicates an almost monotonic inverse relation between earnings and CDS premia. The relation is significant for all but the seven- and ten-year maturities. The same near-monotonic relation holds for both cash flows and accruals. Similar to *ROA*, the relations are significant for all maturities except for the seven- and ten-year maturities.

We also estimate the economic impact of *ROA*, accruals, and cash flows, for the different maturities. We find that the impact of a 1 percent change in *ROA* would change the CDS premium by between 5 to 6 basis points irrespective of the contract maturity. Similar patterns characterize cash flows and accruals. A 1 percent change in cash flows scaled by total assets (accruals scaled by total assets) would change the CDS premium by between 5 to 6 (4 to 5) basis points, irrespective of the maturity.[10]

Overall, the results in Table 3 are consistent with H1 and H4. Earnings, cash flows and accruals are significantly and inversely related to CDS premia for all but the longest CDS maturities, and the relations decline almost monotonically with maturity.

---

[9] Initially, we also included the standard deviation of earnings as an additional measure of accounting information certainty/transparency, but it proved to be insignificant in all regressions.

[10] We obtain economic significance by measuring the dependent variable as a percent (number of basis points divided by 100). This specification is more convenient for estimating the economic impact because the coefficients of the independent variables represent the impact of 1 percent change in the variables on CDS premia (in basis points).

Changes Analysis

The panel data levels analysis assumes *inter alia* that CDS premia are stationary. If CDS premia are nonstationary, then levels regressions will generally yield biased coefficient estimates and biased standard errors. To test for intercept and trend stationarity, we use the KPSS test statistic (Kwiatkowski et al. 1992) for (unbalanced) panel data developed by Hadri and Larsson (2005). We find that stationarity of CDS premia levels can be uniformly rejected for all maturities and for all restructuring clauses. In contrast, stationarity often cannot be rejected for changes in CDS premia. To mitigate concerns about stationarity and omitted correlated variables, we investigate the relation between earnings and CDS rates using a changes analysis.

Table 4 tests H2 regarding quarterly changes in CDS premia. Since CDS instruments trade in an over-the-counter market, it is not possible to see how CDS rates change over time for the same contract. Instead, we measure quarterly changes by comparing two different contracts with the same contractual features for the same reference entity. The existence of multiple contracts from the same reference entity makes such an analysis possible. To the extent that this induces noise into the dependent variable, it should bias against finding meaningful results.[11]

The dependent variable is the quarterly percentage change in the CDS premium as computed from one day after the prior quarter's SEC filing date to one day after the current quarter's SEC filing date.[12] We favor this window because all quarterly financial statement information becomes available to the market on the SEC filing date. Except for the indicator variables, the independent variables in Table 4 are the quarterly changes in the independent variables of the levels analysis (Table 3). Because positive changes in ratings have different implications for credit risk than negative changes, we separate ratings changes into positive and negative components.

The *BASE* column in Panel A of Table 4 provides coefficient estimates for the baseline changes model. The change in the CDS premium is positively and significantly related to the change in leverage ($p = 0.022$, two-tailed). As expected, the change in the CDS premium is positively (negatively) and significantly related to the change in the volatility of returns (the risk-free rate) ($p < 0.010$, two-tailed). The change in the CDS premium is also positively and significantly related to decreases in the credit rating ($p = 0.004$, two-tailed), but is not significantly related to increases in the credit rating ($p = 0.674$, two-tailed). Further, the change in the CDS premium is monotonically related to maturity, but the relation is significant only for higher maturities. Finally, changes in CDS premia are unrelated to restructuring clauses, probably reflecting the fact that restructuring clauses for our sample are sticky over time.

The *ROA* column in Panel A of Table 4 includes changes in *ROA and* the equity stock return in addition to the *BASE* regression explanatory variables. The change in *ROA* coefficient is negative and significant ($p = 0.023$, two-tailed) consistent with H2. The equity return variable is also negative and significant ($p < 0.001$, two-tailed). These results indicate that earnings convey information about firm credit risk beyond the information contained in equity returns. Except for leverage, which remains positive but insignificant at conventional levels, the remaining coefficients are qualitatively similar to the *BASE* regression.

The *CF_ACC* column breaks down the change in earnings into changes in accruals and cash flows. Both the changes in accruals ($p = 0.042$, two-tailed) and cash flows (p

---

[11] Some noise is inevitable since no data are available regarding the parties to the transactions beyond the same reference entity. Thus, the contracts whose rates are compared may involve different contracting parties, and consequently different counterparty risk.

[12] Similar results obtain when the dependent variables is measured as the change in CDS premia.

TABLE 4
Regressions of the Changes in Quarterly CDS Premia (Premice)

Panel A: Regression of the Change in Quarterly CDS Premium on the Change in Earnings, Cash Flows, and Accruals[a]

	BASE	*ROA*	*CF_ACC*
Intercept	−0.015	−0.023	−0.023
	(0.813)	(0.717)	(0.705)
DEL_ROA		−1.693**	
		(0.023)	
DEL_OCF			−2.088**
			(0.024)
DEL_ACC			−1.539**
			(0.042)
DEL_LEV	0.813**	0.407	0.381
	(0.022)	(0.235)	(0.258)
DEL_SPOT	−0.329***	−0.329***	−0.328***
	(0.001)	(0.000)	(0.000)
DEL_SDRET	7.483***	7.195***	7.307***
	(0.000)	(0.000)	(0.000)
POS_DRATE	−0.027	−0.037	−0.038
	(0.674)	(0.490)	(0.485)
NEG_DRATE	0.212***	0.177***	0.177***
	(0.004)	(0.004)	(0.004)
CUMR	−0.012	−0.004	−0.004
	(0.425)	(0.776)	(0.777)
EXR	−0.008	−0.009	−0.009
	(0.391)	(0.338)	(0.341)
D_3	−0.004	−0.006	−0.006
	(0.662)	(0.527)	(0.546)
D_5	0.014	0.011	0.012
	(0.188)	(0.279)	(0.266)
D_7	0.019*	0.014	0.014
	(0.098)	(0.196)	(0.190)
D_10	0.024**	0.020*	0.020*
	(0.040)	(0.082)	(0.080)
EQ_RET		−0.669***	−0.668***
		(0.000)	(0.000)
Adj. R^2	0.29	0.332	0.333
Chi-square	358***	406***	408***
n		7,225 contracts; 362 firms	

Panel B: Changes Regressions with Maturity Interaction[b]

	ROA	*CF_ACC*
Intercept	−0.027	−0.027
	(0.660)	(0.657)
DEL_ROA	−1.152	
	(0.269)	
DEL_ROA_3	−0.757	
	(0.193)	

(*continued on next page*)

TABLE 4 (continued)

	ROA	*CF_ACC*
DEL_ROA_5	−0.613 (0.379)	
DEL_ROA_7	−0.796 (0.309)	
DEL_ROA_10	−0.342 (0.674)	
DEL_OCF		−2.142* (0.087)
DEL_OCF_3		−0.189 (0.788)
DEL_OCF_5		0.169 (0.837)
DEL_OCF_7		−0.133 (0.880)
DEL_OCF_10		0.439 (0.631)
DEL_ACC		−0.833 (0.434)
DEL_ACC_3		−0.815 (0.192)
DEL_ACC_5		−0.796 (0.272)
DEL_ACC_7		−0.994 (0.220)
DEL_ACC_10		−0.651 (0.447)
DEL_LEV	0.406 (0.236)	0.384 (0.252)
DEL_SPOT	−0.329*** (0.000)	−0.330*** (0.000)
DEL_SDRET	7.203*** (0.000)	7.299*** (0.000)
POS_DRATE	−0.037 (0.493)	−0.037 (0.487)
NEG_DRATE	0.177*** (0.004)	0.176*** (0.004)
CUMR	−0.004 (0.770)	−0.004 (0.789)
EXR	−0.009 (0.335)	−0.009 (0.347)
D_2	−0.006 (0.519)	−0.006 (0.553)
D_3	0.011 (0.277)	0.012 (0.254)
D_4	0.014 (0.194)	0.015 (0.177)
D_5	0.020* (0.082)	0.021* (0.072)

(*continued on next page*)

TABLE 4 (continued)

	ROA	*CF_ACC*
EQ_RET	−0.669*** (0.000)	−0.665*** (0.000)
Adj. R^2	0.329	0.330
Chi-square	419***	466***
n	7,225 contracts; 362 firms	

Panel C: Tests for Change in Earnings, Accruals, and Cash Flows across Maturities[c]

	DEL_ROA	*DEL_ACC*	*DEL_OCF*
1-Year CDS	−1.152 (0.269)	−0.833 (0.434)	−2.142* (0.087)
3-Year CDS	−1.909** (0.028)	−1.649* (0.066)	−2.331** (0.023)
5-Year CDS	−1.765** (0.013)	−1.630** (0.024)	−1.973** (0.024)
7-Year CDS	−1.948** (0.012)	−1.827** (0.017)	−2.275** (0.018)
10-Year CDS	−1.494** (0.050)	−1.485** (0.050)	−1.703* (0.073)

---

*, **, *** Indicates statistical significance at the 10 percent, 5 percent, 1 percent levels, respectively.

[a] Panel A shows the regression results of the quarterly change in CDS premia on the change in earnings and the change in accruals and cash flows. The dependent variable is the change in CDS premium during the quarter (*CDS_QTR*). The regressions are estimated using firm random effects with year fixed effects (suppressed). We use robust standard errors clustered by firm. The two-tailed p-values of the coefficients are in parentheses.

[b] Panel B shows the results of the regressions in Panel A above when interacting *DEL_ROA*, *DEL_ACC*, and *DEL_OCF* with the maturity dummies. The *DEL_ROA* coefficients with suffix *i* (*i* = 3,5,7,10) represent the interaction of *DEL_ROA* with *D_i*, an indicator variable equal to 1 if the CDS contract maturity is *i*, and 0 otherwise. Thus, *DEL_ROA_i* represents the incremental effect of *DEL_ROA* on the premium of the CDS contract with maturity *i* over the impact of *DEL_ROA* on the premium of the CDS contract with 1 year to maturity. The *CF_ACC* column replaces *DEL_ROA* with *DEL_ACC* and *DEL_OCF*. To avoid potential multicollinearity and to be able to interpret interactions correctly, *DEL_ROA*, *DEL_ACC*, and *DEL_OCF* are demeaned prior to creating the interaction variables. The regressions are estimated using firm random effects with year fixed effects (suppressed). We use robust standard errors clustered by firm.

[c] Panel C shows the overall coefficients for *DEL_ROA*, *DEL_ACC*, and *DEL_OCF* for the different maturities. For example, the *DEL_ROA* coefficient on the three-year CDS equals the *DEL_ROA* coefficient plus the *DEL_ROA_3* coefficient from Panel B. The two-tailed p-values of the coefficients are in parentheses.

Variable Definitions:

DEL_ACC (*DEL_OCF*) = change in accruals (cash flows);

DEL_SPOT = change in the one-year T-Bill rate;

DEL_SDRET = change in the standard deviation of the firm's stock return during the quarter from the previous quarter;

NEG_DRATE (*POS_DRATE*) = an indicator variable equal 1 if the firm experienced a decrease (increase) in credit rating, and 0 otherwise;

EQ_RET = firm's stock return from the first day after the previous SEC filing date to the first day after the current quarter SEC filing date; and

n = number of observations (*CDS* contracts and firms).

All other variables are as defined in previous tables.

= 0.024, two-tailed) are significant and negative. Again, the accruals changes coefficient is smaller in absolute value than the cash flow changes coefficient, suggesting that accrual earnings are less credible to investors than cash flows. However, the *difference* between the cash flow and accrual earnings coefficients is not significant at conventional levels based on robust clustered standard errors. Using conventional robust standard errors instead, we reject the hypothesis that the coefficients are equal ($p = 0.007$). Further, the coefficient on the change in leverage is significant ($p = 0.001$). The remaining coefficients are qualitatively similar to those obtained in the *BASE* and *ROA* regressions.

Panel B of Table 4 replicates Panel A, inclusive of interaction terms in which maturity is interacted with changes in *ROA*, accruals, and cash flows. Although negatively related to changes in CDS premia, and almost monotonically decreasing in absolute value with maturity, none of the *marginal* interaction terms are significant in the *ROA* column. Similarly, none of the cash flow or accrual changes interaction terms are significant.

Panel C of Table 4 lists the total (as opposed to marginal) coefficients for all maturities. The change in the CDS premium is negatively related to changes in *ROA*, changes in cash flows and changes in accruals, significantly so for all but the one-year maturity. There is no apparent explanation for why the one-year maturity differs from the other maturities. When we replicate the same regressions for only the one-year maturity earnings, accruals, and cash flows (untabulated) we still find that the coefficients on earnings ($p = 0.150$, two-tailed) and accruals ($p = 0.300$, two-tailed) are not statistically significant, while the coefficient on cash flows is significant ($p < 0.050$). Thus, the insignificance of the change in earnings appears to be driven by the lack of significance of accruals, which tend to reverse within one year.

The economic impact of *ROA*, accruals, and cash flows is fairly similar across different maturities. Specifically, a 1 percent change in *ROA* modifies the CDS premium by between 3.5 and 4.5 basis points. Repeating the analysis for accruals and cash flows, we find that a 1 percent change in cash flows (accruals) scaled by total assets alters the CDS premium by about 5 (4) basis points, irrespective of contract maturity.

Event Study Analysis

Table 5 tests H3, the relation between short-window changes in CDS premia in the three days centered on the preliminary earnings announcements. As with our earlier analyses, these changes are measured by comparing two different contracts with the same contractual features for the same reference entity over the short window. Panel A of Table 5 regresses the percentage change of the CDS premium on the earnings surprise, controlling for restructuring clauses and contract maturity. The first column of Panel A shows that there is a significant inverse relation between the earnings surprise and the changes in the CDS premium over this short window ($p = 0.017$, two-tailed).

Following Ball and Brown (1968), a large literature indicates that security returns incorporate the information conveyed by earnings. Thus, it is possible that earnings per se will not provide the CDS market with credit default information beyond that contained in equity markets. Column (2) of Table 5, Panel A controls for cumulative equity returns over the same three-day window. The earnings surprise coefficient remains negative but only marginally significant ($p = 0.091$, two-tailed). The equity return is negative and highly significant ($p < 0.001$, two-tailed).

Column (3) of Table 5, Panel A includes interaction terms of earnings with maturity (denoted by *EAR_SUR* for the one-year maturity and *EAR_SUR_i* for maturities $i = 3,5,7,10$). Results indicate that marginal impact of earnings is negative and significant ($p = 0.013$, two-tailed) for the one-year maturity CDS instruments. Together with Panel B, which shows total earnings coefficients, evidence indicates that the information conveyed by earnings alone (beyond the earnings information contained in equity returns) is concentrated in shorter term maturities (one and three years).

TABLE 5
Changes in CDS Premia around Preliminary Earnings Release Date

Panel A: Regression of the Change in CDS Premium on Earnings Surprise[a]

	1	2	3	4	5
Intercept	−0.007	−0.004	−0.005	−0.004	−0.005
	(0.484)	(0.688)	(0.647)	(0.670)	(0.600)
EQ_RET		−0.313***	−0.313***		
		(0.000)	(0.000)		
RES_RET				−0.319***	−0.319***
				(0.000)	(0.000)
EAR_SUR	−1.400**	−0.811*		−1.392***	
	(0.017)	(0.091)		(0.002)	
EAR_SUR_3			−1.265**		−1.834***
			(0.013)		(0.001)
EAR_SUR_5			0.302		0.288
			(0.700)		(0.711)
EAR_SUR_7			0.642		0.635
			(0.418)		(0.419)
EAR_SUR_10			0.722		0.702
			(0.414)		(0.425)
CUMR			0.445		0.434
			(0.595)		(0.601)
EXR	0.002	0.002	0.002	0.002	0.002
	(0.558)	(0.644)	(0.648)	(0.621)	(0.621)
D_3	0.002	0.002	0.002	0.002	0.002
	(0.336)	(0.443)	(0.441)	(0.434)	(0.431)
D_5	−0.000	−0.000	−0.000	−0.000	−0.000
	(0.913)	(0.873)	(0.861)	(0.873)	(0.862)
D_7	−0.000	−0.000	−0.000	−0.000	−0.000
	(0.924)	(0.930)	(0.924)	(0.924)	(0.918)
D_10	0.002	0.002	0.002	0.002	0.002
	(0.484)	(0.501)	(0.505)	(0.490)	(0.493)
Adj. R^2	0.009	0.043	0.043	0.043	0.043
Chi-square	12.39	62***	74***	62***	74***
n			9,109 contracts; 383 firms		

Panel B: Tests for Earnings Surprise across Maturities[b]

	EQ_RET	*RES_RET*
1-Year CDS	−1.265**	−1.834***
	(0.013)	(0.000)
3-Year CDS	−0.963*	−1.546***
	(0.066)	(0.001)
5-Year CDS	−0.623	−1.199**
	(0.234)	(0.015)
7-Year CDS	−0.543	−1.132*
	(0.620)	(0.054)
10-Year CDS	−0.820	−1.400***
	(0.156)	(0.010)

(*continued on next page*)

TABLE 5 (continued)

*, **, *** Indicates statistical significance at the 10 percent, 5 percent, and 1 percent levels, respectively.

[a] Panel A shows the results of the analysis of the change in CDS premia in the three-day window around the preliminary earnings release date. The dependent variable is the change in CDS premium over the three days centered on the preliminary earnings release date.

[b] Panel B shows the overall coefficients for *EAR_SUR* for the different maturities. For example, the *EAR_SUR* coefficient on the three-year CDS equals to the *EAR_SUR* coefficient plus the *EAR_SUR_3* coefficient from Panel B.

Variable Definitions:

EAR_SUR = earnings surprise (see exact definition in the notes to Table 2);

EQ_RET = equity return during the three-day preliminary earnings window;

RES_RET = residual from the regression of *EQ_RET* on *EAR_SUR*. The *EAR_SUR* coefficients with suffix *i* (i = 3,5,7,10) represent the interaction of *EAR_SUR* with *D_i*, an indicator variable equal 1 if the CDS contract maturity is *i*, and 0 otherwise. Thus, *EAR_SUR_i* represents the incremental effect of *EAR_SUR* on the premium of the CDS contract with maturity *i* over the impact of *EAR_SUR* on the premium of the CDS contract with one year to maturity. To avoid potential multicollinearity and to be able to interpret interactions correctly, *EAR_SUR* is demeaned prior to creating the interaction variables; and

n = number of observations (*CDS* contracts and firms).

The other control variables are defined in previous tables. The regressions are estimated using firm random effects with year fixed effects (suppressed). We use robust standard errors clustered by firm. The two-tailed p-values of the coefficients are in parentheses.

Although it may be that earnings convey little credit risk information beyond equity returns, this does not mean that earnings in total have a marginal impact on CDS pricing since, after all, cumulative equity returns around earnings announcements are driven primarily by earnings information.[13] To see the extent to which earnings per se convey credit risk information over the short window, whether directly or indirectly through equity prices, we regress cumulative equity returns over the short window on earnings. The residual of this regression (denoted *RES_RET*) measures the information contained in equity prices *other than earnings*. Column (4) of Table 5 regresses CDS rates on the earnings surprise and on *RES_RET*. Both the earnings surprise and *RES_RET* coefficients are negative and highly significant ($p = 0.002$ and $p < 0.001$, two-tailed, respectively). Column (5) replicates this regression after interacting the earnings surprise with the different CDS contract maturities. Panel B shows that once earnings are defined all-inclusively to include the earnings information about credit risk conveyed by equity prices, earnings provide significant credit risk information for all CDS maturities.

Tables 6 and 7 test for the nonlinearity of debt holder payoff functions in CDS markets. Table 6 shows the levels, quarterly changes, and short-window results when the firms are divided into high- and low-ratings firms, with a rating of BBB and below defined as low. Panel A shows the levels regression. The column labeled *ROA* regresses CDS premia on earnings decomposed into earnings for high- and low-ratings firms. The coefficient for low-ratings firms is negative and highly significant ($p = 0.005$, two-tailed), whereas the coefficient for high-ratings firms is negative and only significant at the one-tailed level ($p = 0.132$, two-tailed). Moreover, the low-ratings coefficient is larger in absolute value than the high-ratings coefficient, although the difference is not statistically different at conventional levels.[14] The column labeled *CF_ACC* decomposes earnings into cash flows and

---

[13] Although Callen et al. (2006) show that both risk and earnings drive security returns around short-window preliminary earnings announcements, earnings are by far the primary driver.

[14] Again, using a conventional or robust standard error, we would reject equality.

TABLE 6
Nonlinearity of the CDS Premium and Earnings Relation by High/Low Credit Ratings

Panel A: Levels Analysis[a]

	ROA	*CF_ACC*
Intercept	6.238***	6.227***
	(0.000)	(0.000)
ROA_HI	−2.261	
	(0.132)	
ROA_LO	−4.240***	
	(0.005)	
OCF_HI		−2.416
		(0.105)
OCF_LO		−4.590***
		(0.003)
ACC_HI		−2.581*
		(0.098)
ACC_LO		−3.264*
		(0.050)
LEV	1.146***	1.139***
	(0.000)	(0.000)
SPOT	−0.117***	−0.114***
	(0.000)	(0.000)
SD_RET	15.808***	15.857***
	(0.000)	(0.000)
RATE	0.039**	0.042**
	(0.041)	(0.032)
CUMR	0.004	0.005
	(0.706)	(0.640)
EXR	−0.075***	−0.075***
	(0.000)	(0.000)
D_3	0.261***	0.261***
	(0.000)	(0.000)
D_5	0.467***	0.467***
	(0.000)	(0.000)
D_7	0.565***	0.565***
	(0.000)	(0.000)
D_10	0.682***	0.682***
	(0.000)	(0.000)
SIZE	−0.328***	−0.330***
	(0.000)	(0.000)
Adj. R^2	0.529	0.529
Chi-square	2670***	2740***
n	12,220 contracts; 423 firms	

Panel B: Changes in Quarterly CDS Premiums[b]

	ROA	*CF_ACC*
Intercept	−0.014	−0.014
	(0.829)	(0.818)
DEL_ROA_HI	−2.342	
	(0.160)	

(*continued on next page*)

TABLE 6 (continued)

	ROA	*CF_ACC*
DEL_ROA_LO	−1.203*	
	(0.096)	
DEL_OCF_HI		−2.496
		(0.143)
DEL_OCF_LO		−1.598
		(0.159)
DEL_ACC_HI		−2.271
		(0.194)
DEL_ACC_LO		−0.972
		(0.177)
DEL_LEV	0.306	0.287
	(0.388)	(0.412)
DEL_SPOT	−0.304***	−0.303***
	(0.001)	(0.001)
DEL_SDRET	6.628***	6.733***
	(0.000)	(0.000)
NEG_DRATE	−0.031	−0.031
	(0.565)	(0.571)
POS_DRATE	0.178***	0.178***
	(0.003)	(0.003)
CUMR	−0.007	−0.007
	(0.651)	(0.632)
EXR	−0.011	−0.011
	(0.243)	(0.249)
D_2	−0.006	−0.006
	(0.520)	(0.537)
D_3	0.011	0.012
	(0.278)	(0.270)
D_4	0.014	0.014
	(0.204)	(0.198)
D_5	0.020*	0.020*
	(0.097)	(0.095)
EQ_RET	−0.697***	−0.695***
	(0.000)	(0.000)
Adj. R^2	0.333	0.333
Chi-square	410***	427***
n	7,102 contracts; 350 firms	

Panel C: Changes in CDS Premia around Preliminary Earnings Announcements[c]

	1	2
Intercept	−0.004	−0.004
	(0.685)	(0.639)
EQRET	−0.314***	
	(0.000)	
RES_RET		−0.314***
		(0.000)
EAR_SUR_LO	−0.847*	−1.450***
	(0.093)	(0.002)

(*continued on next page*)

TABLE 6 (continued)

	1	2
EAR_SUR_HI	−0.178 (0.919)	−0.781 (0.652)
CUMR	0.002 (0.639)	0.002 (0.639)
EXR	0.002 (0.434)	0.002 (0.434)
D_3	−0.000 (0.876)	−0.000 (0.876)
D_5	−0.000 (0.933)	−0.000 (0.933)
D_7	0.002 (0.501)	0.002 (0.501)
D_10	0.003 (0.278)	0.003 (0.278)
Adj. R^2	0.043	0.043
Chi-square	62***	62**
n	9,109 contacts; 383 firms	

---

*, **, *** Indicates statistical significance at the 10 percent, 5 percent, 1 percent level, respectively.

Table 6 shows the results of the analysis of the impact of earnings on CDS premia conditioned on high/low credit ratings. We define high ratings as BBB+ (Code 10) and up, and low ratings as BBB (Code 11) and below. The regressions are estimated using firm random effects with year fixed effects (suppressed). We use robust standard errors clustered by firm.

[a] Panel A shows the levels analysis.

[b] Panel B shows the quarterly regression analysis.

[c] Panel C shows the analysis of the change in CDS premiums in the three-day window around the preliminary earnings release date.

Variable Definitions:

ROA_HI (*ROA_LO*) = *ROA* if credit rating is high (low) and 0 otherwise; *OCF_HI* (*OCF_LO*) and *ACC_HI* (*ACC_LO*) are defined similarly;
DEL_ROA_HI (*DEL_ROA_LO*) = *DEL_ROA* if credit rating is high (low) and 0 otherwise; *DEL_OCF_HI* (*DEL_OCF_LO*) and *DEL_ACC_HI* (*DEL_ACC_LO*) are defined similarly;
EAR_SUR_HI (*EAR_SUR_LO*) = *EAR_SUR* if credit rating is high (low); and
n = number of observations.

All other variables are defined in Tables 3, 4, and 5.

accruals for high-/low-ratings firms. Coefficients other than cash flow high ratings are negative and significant at conventional two-tailed levels. The coefficient for cash flow high-ratings firms is negative and marginally significant ($p = 0.105$, two-tailed). Again, the low-ratings coefficients are larger than the high-ratings coefficients in absolute values for both cash flows and accruals, but the differences are not significant at conventional levels. Note that the cash flow low-ratings coefficient is larger in absolute value than the accrual low-ratings coefficient (p-value = 0.09, two-tailed) consistent with accruals conveying less information about credit risk than cash flows.

Nonlinearity is also apparent when examining the economic impact of earnings, cash flows, and accruals for high-/low-ratings firms. Earnings information has much greater economic impact on CDS for low-ratings firms. Specifically, a 1 percent change in *ROA*, cash flows, or accruals of low-ratings firms changes CDS premia by close to 10 basis points (approximately 20 percent), whereas a similar change in any of these variables for high-ratings firms has almost no economically or statistically significant impact on CDS premia.

Panel B of Table 6 shows the quarterly changes regression. The column labeled *ROA* shows the regression of the quarterly change in CDS premia on the change in earnings, decomposed into changes in earnings for high- and low-ratings firms. The low-ratings earnings coefficient is significant and negative (p = 0.096, two tailed). The high-ratings earnings coefficient is significant and negative only at the one-tailed level (p = 0.160, two-tailed). The further breakdown of earnings into accruals and cash flows yields negative but insignificant coefficients. Similarly, the changes analysis indicates no significant difference between high-/low-ratings firms regarding earnings and its components.

Column 1 of Table 6, Panel C shows the short-window analysis, with the earnings surprise decomposed into high- and low-ratings firms controlling for stock returns. The high-ratings earnings surprise coefficient is not significant, whereas the low-ratings coefficient is negative and significant (p = 0.093, two-tailed). The equity return is negative and highly significant. Column 2 shows the results when we replace equity return with the residual from the regression of stock return on the earnings surprise. The results are virtually identical—the earnings surprise coefficient is significant only for low credit ratings firms.

Table 7, Panel A shows the levels regression for firms with *ROA* above and below the industry median. We define an indicator variable (*LOWEARN*) equal to 1 if earnings are below the median, and 0 otherwise. In addition to the separate regressors *ROA* and *LOWEARN*, the independent variables also include their interaction (*LOWEARN* * *ROA*). The column labeled *ROA* shows that *LOWEARN* is positive and highly significant (p = 0.006, two-tailed) consistent with CDS premia being higher for below median earnings firms. More importantly, although the above median earnings coefficient is not significant, Panel B indicates that the below median earnings coefficient of −3.119 (= 1.234 − 4.353) is negative and significant (p = 0.062, two-tailed). The asymmetric reaction of CDS premia to earnings is consistent with debt holders being more concerned with poor earnings. The column labeled *CF_ACC* in Panel A and the analysis in Panel B show similar results for cash flows and accruals. More specifically, the cash flow coefficient of −3.453 (= 0.978 − 4.431) is negative and significant (p = 0.051, two-tailed) for below median earnings firms and insignificant for above median earnings firms. The accruals coefficient is negative and significant only at the one-tailed level for above median earnings firms (p = 0.111, two-tailed). As with high-/low-ratings firms, we find that the economic impact of *ROA*, cash flows, and accruals is economically and statistically significant for firms with earnings below the median. A 1 percent change in *ROA*, cash flows, or accruals for firms that have *ROA* below the median would change the CDS premium by 7, 7.2, and 6.5 basis points, respectively. The impact of a change in either variable for firms with ROA above the median is not economically or statistically significant.

Panels C and D of Table 7 show a similar but stronger pattern of results for the changes regression. Specifically, we find that the change in *ROA* (accruals, cash flow) is significantly and negatively related to the change in CDS premia only for firms with *ROA* (accruals, cash flow) below the median. Consistent with the levels analysis, the changes analysis also indicates that changes in *ROA*, cash flows, and accruals have a significant economic impact on CDS premia only for firms with *ROA* below the median. For these firms, a 1 percent change in *ROA*, cash flows, or accruals alters the CDS premium by 4.4, 6, and 4 basis points, respectively. In contrast, untabulated results indicate that the short-window event study analysis does not confirm the asymmetry findings of the levels and changes analyses, consistent with Datta and Dhillon (1993) and in contradiction to those of Easton et al. (2007) and DeFond and Zhang (2008) in the corporate bond market.

To test the robustness of our results, we undertook a number of further (untabulated) analyses. We examine whether results are sensitive to the inclusion of the standard deviation

TABLE 7
Nonlinearity of the CDS Premium and Earnings Relation by ROA Above/Below the Industry Median

Panel A: Levels Analysis[a]

	ROA	*CF_ACC*
Intercept	6.177***	6.192***
	(0.000)	(0.000)
LOWEARN	0.150***	0.151***
	(0.006)	(0.005)
ROA	1.234	
	(0.528)	
*LOWEARN * ROA*	−4.353	
	(0.104)	
OCF		0.978
		(0.611)
*LOWEARN * OCF*		−4.431
		(0.107)
ACC		1.792
		(0.402)
*LOWEARN * ACC*		−4.403
		(0.110)
LEV	1.103***	1.094***
	(0.000)	(0.000)
SPOT	−0.116***	−0.113***
	(0.000)	(0.000)
SD_RET	15.965***	15.906***
	(0.000)	(0.000)
RATE	0.037**	0.037**
	(0.050)	(0.048)
CUMR	0.005	0.005
	(0.651)	(0.643)
EXR	−0.075***	−0.075***
	(0.000)	(0.000)
D_3	0.261***	0.261***
	(0.000)	(0.000)
D_5	0.467***	0.467***
	(0.000)	(0.000)
D_7	0.565***	0.565***
	(0.000)	(0.000)
D_10	0.683***	0.683***
	(0.000)	(0.000)
SIZE	−0.331***	−0.331***
	(0.000)	(0.000)
Adj. R^2	0.534	0.534
Chi-square	2626***	2659***
n	12,220 contracts; 423 firms	

(*continued on next page*)

TABLE 7 (continued)

Panel B: Tests for Earnings below the Median[b]

*DEL_ROA + LOWEARN * DEL_ROA*	−3.119* (0.062)	
*DEL_OCF + LOWEARN * DEL_OCF*		−3.453* (0.051)
*DEL_ACC + LOWEARN * DEL_ACC*		−2.611 (0.111)

Panel C: Changes in Quarterly CDS Premiums[c]

	ROA	*CF_ACC*
Intercept	−0.033 (0.595)	−0.035 (0.564)
LOWEARN	0.019 (0.520)	0.021 (0.469)
DEL_ROA	−0.576 (0.683)	
*LOWEARN * DEL_ROA*	−1.459 (0.379)	
DEL_OCF		−0.638 (0.664)
*LOWEARN * DEL_OCF*		−2.356 (0.224)
DEL_ACC		−0.496 (0.743)
*LOWEARN * DEL_ACC*		−1.300 (0.456)
DEL_LEV	0.399 (0.243)	0.374 (0.259)
DEL_SPOT	−0.331*** (0.000)	−0.325*** (0.000)
DEL_SDRET	7.282*** (0.000)	7.369*** (0.000)
POS_DRATE	−0.041 (0.447)	−0.042 (0.441)
NEG_DRATE	0.178*** (0.004)	0.177*** (0.004)
CUMR	−0.005 (0.747)	−0.005 (0.728)
EXR	−0.008 (0.352)	−0.008 (0.384)
D_3	−0.006 (0.546)	−0.005 (0.572)
D_5	0.012 (0.270)	0.012 (0.255)
D_7	0.014 (0.197)	0.015 (0.187)
D_10	0.020* (0.085)	0.020* (0.081)

(*continued on next page*)

TABLE 7 (continued)

	ROA	***CF_ACC***
EQRET	−0.669*** (0.000)	−0.662*** (0.000)
Adj. R^2	0.334	0.335
Chi-square	409***	412***
n	7,225 contracts; 362 firms	
Panel D: Tests for Earnings below the Median[d]		
*DEL_ROA + LOWEARN * DEL_ROA*	−2.036** (0.017)	
*DEL_OCF + LOWEARN * DEL_OCF*		−2.994** (0.026)
*DEL_ACC + LOWEARN * DEL_ACC*		−1.796** (0.034)

*, **, *** Indicates statistical significance at the 10 percent, 5 percent, 1 percent levels, respectively. The regressions are estimated using firm random effects with year fixed effects (suppressed). We use robust standard errors clustered by firm. The two-tailed p-values of the coefficients are in parentheses.

[a] Panel A shows the results of the analysis of the impact of earnings on CDS premia conditioned on *ROA* below/above the three-digit industry median *ROA*.

[b] Panel B shows the overall coefficients for *ROA*, *ACC*, and *OCF* for low *ROA* firms. For example, the coefficient on *ROA* below the median equals the *ROA* coefficient plus the *LOWEARN * ROA* coefficient.

[c] Panel C shows the results of the analysis of the impact of earnings changes on changes in CDS premia conditioned on *ROA* below/above the three-digit industry median *ROA*. *LOWEARN * DEL_ROA*, *LOWEARN * DEL_ACC*, and *LOWEARN * DEL_OCF* are the interaction variables of *LOWEARN* with *DEL_ROA*, *DEL_ACC*, and *DEL_OCF*, respectively.

[d] Panel D shows the overall coefficients for *DEL_ROA*, *DEL_ACC*, and *DEL_OCF* for low *ROA* firms. For example, the coefficient on *DEL_ROA* below the median equals the *DEL_ROA* coefficient plus the *LOWEARN * DEL_ROA* coefficient.

Variable Definitions:

LOWEARN = an indicator variable with 1 if *ROA* is below the industry median *ROA*, and 0 otherwise; *LOWEARN * ROA*, *LOWEARN * ACC*, and *LOWEARN * OCF* are the interaction variables of *LOWEARN* with *ROA*, *ACC* and *OCF*, respectively; and

n = number of observations.

All other variables defined in Tables 3 and 4.

of *ROA* and the market-to-book ratio. In the levels analysis, we include stock returns and the return on the S&P 500 index as proxies for changes in the economic condition of the company and the economy in general. We also control for the slope of the yield curve, defined as the difference between the ten-year and one-year Treasury Bill rates, which serves as a proxy for expectations about future interest rate changes. In the changes regression of Table 3, we include the return on the S&P 500 and the level of earnings among the independent variables and also control for size. We further interact equity returns and all other control variables with maturities. We also redefine leverage to be long-term debt to total assets and redefine the earnings surprise to be standardized unexpected earnings (or its rank), estimated as earnings for quarter t minus earnings for quarter $t-4$ less a drift term scaled by the standard deviation of earnings surprises in the previous eight quarters. The results are robust to these changes.

VII. CONCLUSION

This study examines the relation between CDS and accounting information. Notwithstanding the role of CDS in risk management, overexposure to CDS is seen as one of the drivers of the recent financial crisis. The sharp increase in CDS rates are believed to have played a role in the collapse of Lehman Brothers Holdings Inc. They also likely helped push AIG to seek shelter in the arms of the government (Scannell et al. 2008). At the time this study is written, calls for regulation of CDS have ramped up as these financial instruments have been identified as one of the culprits of the current market stress.

The phenomenal growth in CDS market together with the recent financial crisis demonstrate the importance of swaps to the economy, and consequently the need to further study how CDS prices are determined. Our study contributes by showing that earnings levels (changes) are significantly and negatively correlated with the level of (changes in) CDS premia, consistent with earnings conveying information about default risk. Importantly, the impact of earnings on CDS premia is economically significant. At the median, a 1 percent increase in earnings (normalized by total assets) reduces the CDS premium by 9 percent in the levels regression and by 5 percent in the changes regression. Other determinants of CDS premia are also significant and in the directions predicted by the theory.

REFERENCES

Abid, F., and N. Naifar. 2006. The determinants of credit default swap rates: An explanatory study. *International Journal of Theoretical and Applied Finance* 9 (1): 23–42.

Aiken, L. S., and S. G. West. 1991. *Multiple Regression*. Beverly Hills, CA: Sage Publishing.

Aunon-Nerin, D., D. Cossin, T. Hricko, and Z. Huang. 2002. Exploring for the determinants of credit risk in credit default swap transaction data: Is fixed-income markets information sufficient to evaluate credit risk? Research Paper No. 65. Geneva, Switzerland: FAME.

Ball, R., and P. Brown. 1968. An empirical evaluation of accounting income numbers. *Journal of Accounting Research* 6 (2): 159–178.

Barth, M. E., D. Cram, and K. Nelson. 2001. Accruals and the prediction of future cash flows. *The Accounting Review* 76 (1): 27–58.

Batta, G. 2006. Credit derivatives and financial information. Working paper, University of Southern California.

Benkert, C. 2004. Explaining credit default swap premia. *Journal of Futures Markets* 24 (1): 71–92.

Berndt, A. R. , A. Jarrow, and C. Kang. 2006. Restructuring risk in credit default swaps: An empirical analysis. *Stochastic Processes and Their Applications* 117 (11): 1724–1729.

Blanco, R., S. Brennan, and I. W. Marsh. 2005. An empirical analysis of the dynamic relation between investment-grade bonds and credit default swaps. *The Journal of Finance* 60 (5): 2255–2281.

Callen, J. L., J. Livnat, and D. Segal. 2006. The value relevance and information environment of SEC filings: A variance decomposition analysis. *The Accounting Review* 81 (5): 1017–1043.

Cheng, C. S. A., and W. B. Thomas. 2006. Evidence of the abnormal accrual anomaly incremental to operating cash flows. *The Accounting Review* 81 (5): 1151–1167.

Collin-Dufresne, P., R. S. Goldstein, and S. Martin. 2001. The determinants of credit spread changes. *The Journal of Finance* 56 (6): 2177–2207.

Daniels, K. N., and M. S. Jensen. 2005. The effect of credit ratings on credit default swap spreads and credit spreads. *Journal of Fixed Income* 15 (3): 16–33.

Das, S. R. 1995. Credit risk derivatives. *Journal of Derivatives* 2 (Spring): 7–23.

———, and R. Sundaram. 2000. A discrete-time approach to arbitrage-free pricing of credit derivatives. *Management Science* 46 (1): 46–62.

———, P. Hanouna, and A. Sarin. 2009. Accounting-based versus market-based cross-sectional models of CDS spreads. *Journal of Banking and Finance* 33 (4): 719–730.

Datta, S., and U. S. Dhillon. 1993. Bond and stock market response to unexpected earnings announcements. *Journal of Financial and Quantitative Analysis* 28 (4): 565–577.

Dechow, P. 1994. Accounting earnings and cash flows as measures of firm performance: The role of accounting accruals. *Journal of Accounting and Economics* 18 (1): 3–42.

———, S. Kothari, and R. Watts. 1998. The relation between earnings and cash flows. *Journal of Accounting and Economics* 25 (2): 133–168.

DeFond, M., and J. Zhang. 2008. The information content of earnings surprises in the bond market. Working paper, University of Southern California.

Duffie, D. 1999. Credit swap valuation. *Financial Analysts Journal* 55 (1): 73–87.

———, and D. Lando. 2001. Term structures of credit spreads with incomplete accounting information. *Econometrica* 69 (3): 633–644.

Easton, P. D., S. J. Monahan, and F. Vasvari. 2007. Some initial evidence on the role of accounting earnings in the bond market. Working paper, Notre Dame University, INSEAD, London Business School.

Elton, E. J., M. J. Gruber, D. Agrawal, and C. Mann. 2001. Explaining the rate of spread on corporate bonds. *The Journal of Finance* 56 (1): 247–277.

Ericsson, J., J. Reneby, and H. Wang. 2006. Can structural models price default risk? evidence from the bond and credit derivative markets. Working paper, McGill University and the University of Stockholm.

———, K. Jacobs, and R. Oviedo. 2009. The determinants of credit default swap premia. *Journal of Financial and Quantitative Analysis* 44 (1): 133–154.

Finger, C. 1994. The ability of earnings to predict earnings and cash flows. *Journal of Accounting Research* 32 (2): 210–223.

Hadri, K., and R. Larsson. 2005. Testing for stationarity in heterogeneous panel data where the time dimension is finite. *The Econometrics Journal* 8 (1): 55–69.

Houweling, P., and T. Vorst. 2005. Pricing default swaps: Empirical evidence. *Journal of International Money and Finance* 24 (8): 1200–1225.

Huang, M., and J. Huang. 2002. How much of the corporate-treasury yield spread is due to credit risk? Working paper, Stanford University.

Hull, J., and A. White. 2000a. Valuing credit default swaps, I: No counterparty default risk. *Journal of Derivatives* 8 (1): 29–40.

———, and ———. 2000b. Valuing credit default swaps, II: Including counterparty default risk. *Journal of Derivatives* 8 (3): 29–40.

Kwiatkowski, D., P. C. B. Phillips, P. Schmidt, and Y. Shin. 1992. Testing the null hypothesis of stationarity against the alternative of a unit root. *Journal of Econometrics* 54 (1–3): 91–115.

Livnat, J., and G. Lopez-Espinosa. 2008. Quarterly accruals or cash flows in portfolio construction. *Financial Analysts Journal* 64 (3): 67–79.

Longstaff, F. A., S. Mithal, and E. Neis. 2005. Corporate yield spreads: Default risk or liquidity? New evidence from the credit default swap market. *The Journal of Finance* 60 (5): 2213–2253.

Merton, R. C. 1974. On the pricing of corporate debt: The risk structure of interest rates. *The Journal of Finance* 29 (2): 449–470.

Nissim, D., and S. Penman. 2001. Ratio analysis and equity valuation: From research to practice. *Review of Accounting Studies* 6 (1): 109–154.

Packer, F., and H. Zhu. 2005. Contractual terms and CDS pricing. *BIS Quarterly Review* (March): 89–99.

Petersen, M. A. 2009. Estimating standard errors in finance panel data sets: Comparing approaches. *Review of Financial Studies* 22 (1): 435–480.

Ryan, S. 2007. *Financial Instruments & Institutions: Accounting and Disclosure Rules*. 2nd edition. Hoboken, NJ: J. Wiley & Sons.

Scannell, K., P. Dvorak, J. S. Lublin, and E. Williamson. 2008. Rescue plan stirs calls for deeper regulation. *Wall Street Journal* (September 24).

Sloan, R. 1996. Do stock prices fully reflect information in accruals and cash flows about future earnings? *The Accounting Review* 71 (3): 289–315.

Zhu, H. 2006. An empirical comparison of credit spreads between the bond market and the credit default swap market. *Journal of Financial Services Research* 29 (3): 211–235.

THE ACCOUNTING REVIEW
Vol. 84, No. 5
2009
pp. 1395–1428

American Accounting Association
DOI: 10.2308/accr.2009.84.5.1395

Rules Rather than Discretion in Audit Standards: Going-Concern Opinions in Belgium

Joseph V. Carcello
The University of Tennessee

Ann Vanstraelen
Maastricht University and University of Antwerp

Michael Willenborg
University of Connecticut

ABSTRACT: We study going-concern (GC) reporting in Belgium to examine the effects associated with a shift toward rules-based audit standards. Beginning in 2000, a major revision in Belgian GC audit standards took effect. Among its changes, auditors must ascertain whether their clients are in compliance with two "financial-juridical criteria" for board of directors' GC disclosures. In a study of a sample of private Belgian companies, we report two major findings. First, there is a decrease in auditor Type II errors, particularly by non-Big 6/5 auditors for their clients that fail both criteria. Second, there is an increase in Type I errors, again particularly for companies that fail both criteria. We also conduct an *ex post* analysis of the decrease in Type II errors and the increase in Type I errors. Our findings suggest the standard engenders both favorable and unfavorable effects, the net of which depends on the priorities assigned to the affected parties (creditors, auditors, companies, and employees).

Keywords: *audit standards; auditing; going-concern opinions; Type I and Type II errors.*

Data Availability: *Data are available upon request.*

We thank Allen Blay, Sophie Audousset-Coulier, Larry Brown, Brian Cadman, Ann Jorissen, Tim Louwers, Steven Maijoor, Steve Matsunaga, Bill Messier, Jane Mutchler, Mark Nelson, Terry O'Keefe, Sundaresh Ramnath, Sarah Rice, Mike Stein, Hervé Stolowy, Stan Veliotis, Joe Weber, Mike Wilkins, Pete Wilson, Arnie Wright, and participants at AAA Auditing Meeting in Huntington Beach, AAA Annual Meeting in Anaheim, Boston College, College of William & Mary, EAA Conference in Göteborg, Georgia State University, HEC Paris, Michigan State University, Texas A&M University, University of Kentucky, and University of Oregon. We owe special thanks to two anonymous referees and Bill Kinney (editor) for helpful comments and suggestions, and to Mike Borth for research assistance. Professor Willenborg thanks the University of Connecticut CIBER.

Editor's note: Accepted by Steven Kachelmeier, with thanks to Bill Kinney for serving as editor on previous versions.

Submitted: August 2006
Accepted: January 2009
Published Online: September 2009

I. INTRODUCTION

Assuming standard-setters endorse the premise that auditors modify their opinions due to uncertainty about a company's going-concern (GC) status, should professional guidance refer to pre-announced rules or should it be discretionary in nature, so as to allow auditors a latitude of choice to tailor their decisions to prevailing circumstances? While, at least since Friedman (1948), economists recommend that policymakers follow *rules rather than discretion*, audit standard-setters usually promulgate a substantial degree of discretion regarding GC judgments.[1]

We study the relation between the nature of GC audit standards, in terms of the extent to which they promulgate discretion or rules-based guidance, and auditor reporting for stressed companies. Our context is Belgium where, in 2000, an audit standard for clients experiencing stress became effective. In contrast to its predecessor, a non-binding "circular letter" allowing for considerable discretion, the standard is a major increase in responsibilities, prominent among which auditors must ensure that a company in financial distress complies with certain rules-based, legal requirements for board of director disclosures. Per a 1997 change to Belgian Company Law, if a company has negative retained earnings in the current year or negative net income for both the current and prior years, the board must justify the use of valuation rules that assume a company's GC status. The GC audit standard specifically references these legal requirements for board GC disclosures. If a client meets, or threatens to meet, either of these financial-juridical criteria, the auditor must remind the board of its responsibility to provide GC disclosures and then opine accordingly. As such, because these criteria influence the board's GC disclosure decision and, in turn, the auditor's GC opinion decision, the standard represents a shift toward rules rather than discretion. Put another way, because the standard expressly requires auditors to remind the board of its rules-based GC disclosure responsibilities, there is now an implied minimum for auditors to look at to assess the GC status of their client. Therefore, while Belgian auditors still have discretion *vis-à-vis* their opinion decision, the audit standard reduces the extent of the discretion prevailing during the previous discretionary (circular letter) regime.[2]

Among the reasons the economics literature advocates rules rather than discretion is that rules, given a credible commitment to enforce them (Kydland and Prescott 1977), can restrain a tendency toward leniency stemming from a decision-maker's concern for self-interest. Boot and Thakor (1993) formalize this insight in a model of a regulator's bank closure decision. The regulator seeks to maximize a weighted objective function of societal welfare and personal gain from attaining/protecting its reputation as high quality. Because of this latter component, the researchers show that the regulator's equilibrium closure policy is too lax and allows a bank to continue operations when it otherwise should not. Boot and Thakor (1993, 211) conclude that "[a] simple way to minimize closure-policy distortions is to reduce regulatory discretion by stipulating a minimum (positive) amount of bank book capital needed to avoid closure ... Indeed, we advocate rules rather than discretion."

---

[1] For example, per U.S. audit standards on GC reporting, whereas "substantial doubt" regarding an entity's ability to continue as a going concern is the "trigger point" for report modification, SAS No. 59 intentionally does not provide a definition of either of these constructs (Carmichael and Pany 1993). As for Belgium, because prior to 2000 there was no standard on GC reporting whatsoever, the auditor's decision was arguably even more discretionary.

[2] The rules-based nature of the standard is evident in both the standard itself and in the opinion of practitioners. Per the audit standard, "[t]he requirement for the Board of Directors to justify the application of valuation rules in the assumption of continuity ... should not be considered equal to a situation of discontinuity ... this financial-juridical criterion merely constitutes an incitement for reflection on the possibility of the company to continue its operations"; and "[t]here is no exception on the justification requirement once one of the two criteria is fulfilled. In this way, justification of continuity is required in case of continuous loss even when the balance sheet structure is perfectly healthy, because a lack of profitability can signal an evolution which could become dangerous for a company. Even when a company successfully executes a reorganization plan, but the loss remains on the balance sheet, it is required to justify the possibility of the company to continue its operations." In addition, in an article in a newsletter published by the Belgian Institute of Registered Auditors, a BDO partner acknowledges the importance of the criteria as "the going concern opinion decision does not only relate to a juridical responsibility of the auditor, but also to a social responsibility to timely signal significant continuity problems" Wilmots (2004).

While such reputation motives are likely present, auditors also have the incentive to be lenient simply because they "may be reluctant to disclose perceived [going-concern] problems, fearing the loss of the client should the firm continue in operation" (Kida 1980, 507). And, following DeAngelo's (1981) logic, if as the relative size of a client increases auditors find it harder to maintain independence, then their fear of client loss is increasing in client size. Consistent with this reasoning, extant (including Belgian) literature consistently finds that the presence of a GC opinion for a distressed, soon-to-be-bankrupt company varies inversely with client size.

Constraining the influence of self-interest, there were two factors prior to the adoption of the GC standard that made the "costs of acquiescence" (Coffee 2004) for Belgian auditors arguably rather low. First, litigation risk in Belgium is such that auditors have little fear of a lawsuit following the bankruptcy of a client with a clean opinion. Also, as we mention above, because the only GC audit guidance was in the form of a non-binding circular letter from the Belgian Institute of Registered Auditors, for this time period there was literally no formal GC audit standard to which auditors could be held professionally accountable.[3]

However, following the adoption of the GC standard, the costs of auditor acquiescence should increase because the penalty for rendering an incorrect clean opinion now involves the oversight of the Belgian Institute of Registered Auditors.[4] Given this development, we show that GC rates should increase, per the model in Matsumura et al. (1997). As such, both the frequency of Type I errors (incorrect GCs) should increase and the frequency of Type II errors (incorrect clean opinions) should decrease. In addition, following Boot and Thakor (1993), if Belgium's shift toward a rules-based standard restrains leniency stemming from auditor self-interest, then the decrease in Type II errors should be especially acute for larger companies. Overall, the effects of the decision by Belgian standard-setters to emphasize rules rather than discretion becomes an empirical issue that involves quantifying these likely changes in Type I and Type II errors.

---

[3] Overall, in Belgium, the objective of an audit is to form an opinion about whether the financial statements give a "true and fair view" of shareholders' equity and financial position at the balance sheet date and whether the results accord with relevant legal and administrative requirements. As such, auditors play both reporting and compliance roles, the latter being especially relevant for the private companies that dominate the economy and, per the 4th EU Directive, many of which must undergo audit. Private companies that meet certain criteria must undergo an audit by a member of the Belgian Institute of Registered Auditors. At present, limited liability companies must be audited if they have 100 or more employees, or they satisfy two or more of the following criteria: (1) 50 or more employees; (2) annual turnover (i.e., sales) exclusive of VAT, is equal to or greater than 7.3 million Euros; (3) assets are equal to or greater than 3.65 million Euros. The Belgian Institute of Registered Auditors is a public institute under the authority of the Ministry of Economic Affairs and is assisted by the High Council for Auditing and Accountancy.

[4] The Belgian Institute of Registered Auditors monitors auditor GC reporting, and extracts of their disciplinary decisions reference auditors' failure to mention GC problems and sanctions that range from warnings to suspensions of one month up to one year. For example, per Case 0217/2003/F: "The Institute blames X [publicly available source document omits auditor name] to have issued in its function of statutory auditor of Y [publicly available source document omits company name] an unqualified opinion on the financial statements of 31/12/2000, while at this date, net assets of the company were negative as well as the net result of the year, and that the report of the Administrative Council of March 2001 does not provide the required information by ... Company Law justifying the valuation rules in case of an accumulated loss or 2 subsequent years of a bottom-line loss. In particular, the Institute blames X to have certified without qualification or explanatory paragraph with respect to the application of the valuation rules concerning the going concern of the company. ... The Disciplinary Committee well understands the concern and hesitation of X to issue a more appropriate opinion for a company in difficulties as it is delicate to issue to company Y that is determined to continue its operations. However, the Disciplinary Committee cannot agree with the chosen solution of giving an unqualified opinion without a qualification or explanatory paragraph."

We collect a sample of stressed, private companies filing bankruptcy in the two years prior to and the two years after the change in Belgian audit standards. We also collect a sample of comparably stressed, private companies in the same industries that did not file bankruptcy during these time periods. We study private companies for practical reasons, as they constitute more than 99 percent of Belgian companies subject to audit (Willekens and Gaeremynck 2005), and because the audit opinion for private companies is arguably more important than that for public companies, as private companies do not disseminate many (if any) voluntary disclosures and because some conclude their financial reporting is of lower quality (Chaney et al. 2004; Ball and Shivakumar 2005, 2008). As such, the audit opinion is likely a major piece of information for stakeholders (e.g., workers and suppliers negotiating labor and credit terms, respectively).

For both of these samples, we find that the change to a rules-based standard is associated with a significant increase in GC opinions. For bankrupt companies, the increase in GCs indicates a decrease in Type II errors; especially for non-Big 6/5 auditors who frequently render a GC opinion to a stressed client that is in violation of both financial-juridical criteria (25 of 29, or 86 percent). For non-bankrupt companies, the increase in GCs indicates an increase in Type I errors; again, especially involving companies that violate both financial-juridical criteria. Overall, in terms of marginal effects, our findings suggest Belgium's standard manifests as a 33 percent decrease in Type II errors and a 13 percent increase in Type I errors.

We then focus on GC reporting for bankrupt and non-bankrupt companies by estimating separate regressions for the pre- and post-standard periods. In the post-standard period, the presence of a GC is no longer inversely related to company size. That is, after the standard is put in place, larger stressed companies are no longer less likely to have a GC opinion, *ceteris paribus*. This finding contrasts with essentially all extant literature, including Belgian research (Vanstraelen 2002), that documents a "size effect," consistent with the view that auditor self-interest influences GC decisions. Our findings suggest that a rules-based standard mitigates this self-interest, in that we find no evidence consistent with auditor leniency toward larger clients in the post-standard period.

We also conduct an *ex post* analysis of the decrease in Type II errors and the increase in Type I errors. Findings suggest that the standard engenders favorable effects, notably that creditors of soon-to-be-bankrupt companies receive better early warning; as well as unfavorable effects. For instance, for companies that receive a GC but do not file bankruptcy the next year, there is some evidence of an increase in auditor switching and employee attrition. Whether the benefits outweigh costs, however, depends on how one prioritizes the perspectives of creditors, companies, employees, and auditors (e.g., from a creditor perspective, the effects of the GC standard seem beneficial).

Studies on the effect of changes in standards on auditor GC reporting behavior in the U.S. produce mixed results (Carcello et al. 1995, 1997; Holder-Webb and Wilkins 2000). However, studies in China (DeFond et al. 1999) and in Germany (Ashbaugh and Gassen 2007) report a significant increase in the presence of GC opinions following a change in audit standards. We also extend the literature on the effects of auditor incentives on GC reporting (e.g., McKeown et al. 1991), as we find that after Belgium's shift toward a rules-based standard, the inverse relation between company size and the presence of a GC opinion is no longer evident. In addition, we contribute to the audit literature on private companies (e.g., Chaney et al. 2004) by examining the effect of a change in audit standards on the determinants of GC opinions for a sample of private companies. Finally, because we study the effects of a change in GC audit standards in a low-litigation environment for a sample of small, private companies, our study may have implications for audit standard-setters in comparable settings.

II. BACKGROUND AND MOTIVATION

GC Reporting and the Benefits and Costs of Auditor Acquiescence in Belgium

If an auditor renders a GC opinion and it proves *ex post* inappropriate (Type I error), then the auditor risks exposure to client loss. Fearful of this consequence, and intent on retaining future economic benefits, auditors have the incentive to acquiesce to management's position that there is no substantial doubt regarding GC status and to render a clean opinion. In Belgium, because of restrictions on auditor consulting, employment opportunities, and career mobility (Maijoor et al. 1998), such benefits primarily relate to audit fees. Because larger companies pay larger audit fees, the benefits of auditors' acquiescence are increasing in client size. Consistent with this reasoning, Vanstraelen (2002), and many U.S. researchers (e.g., Mutchler 1985, 1986; Krishnan and Krishnan 1996; Mutchler et al. 1997; Louwers 1998), find that the presence of a GC opinion on the financial statements of a stressed Belgian company varies inversely with company size. Among the U.S. studies, McKeown et al. (1991, 11) find that this relation persists after controlling for a relation between company size and bankruptcy, which "raises the issue as to whether auditors' independence is influenced by the size of the client."[5]

If an auditor renders a clean opinion and it proves *ex post* inappropriate (Type II error), then, depending on the setting, the auditor risks exposure to litigation costs. For example, in the high-litigation U.S. context, literature supports a relation among GC reporting for stressed clients, company bankruptcy, and auditor litigation risk (St. Pierre and Anderson 1984; Palmrose 1987; Carcello and Palmrose 1994; Geiger and Raghunandan 2001; Francis and Krishnan 2002). In contrast, the risk of auditor litigation in Belgium is rather benign.[6] Moreover, as is typical of a continental European country, Belgian law stems from (French) civil tradition that is based on statutes and codes, banks are major providers of capital (Mueller et al. 1994), investor rights are low (LaPorta et al. 1998) and, before the new GC standard, "auditing guidelines on going concern assessment ... are rather subjective, general, and ambiguous" (Vanstraelen 1999, 44). Collectively, these factors decrease an investor's demand for auditing and provide auditors with substantial discretion regarding GC reporting.[7] Overall, prior to the GC standard the costs of auditor acquiescence in Belgium are arguably rather low.

---

[5] Prior papers also raise this issue (e.g., see Goldman and Barlev [1974, 709] "the potential loss of the audit fee may be a much greater cause for conflict between professional integrity and self-interest"; and DeAngelo [1981]).

[6] Aspects of Belgian law that lower litigation risk for auditors include: no absolute right to a jury trial (judges usually hear cases involving technical accounting evidence); an absence of contingent legal fees; a lack of class action lawsuits; and less liberal discovery rules. Per Gaeremynck and Willekens (2003), "[a]nother notable difference between the American and Belgian audit environment is the lack of auditor litigation in Belgium. Only eight cases against external auditors made it to court since the creation of the Belgian Kingdom in 1831 (Aerts 2002). In litigious environments the threat of litigation works as a deterrent against below-standard quality."

[7] Conversely, the demand for audited information by other stakeholders (e.g., creditors) in Belgium is likely high. For example, Ashbaugh and Warfield (2003) conclude that audits serve a corporate governance role in Germany, a country with many institutional characteristics that are similar to those in Belgium.

Belgium's GC Audit Standard

In 2000, the Belgian Institute of Registered Auditors increased auditor GC reporting responsibilities by issuing a formal standard for audits of clients experiencing stress. In contrast to its predecessor, the standard represents a major change in auditor responsibilities and a shift toward rules-based policymaking. That is, prior to its issuance, there was no audit standard relating to GC reporting whatsoever (the only guidance being a non-binding circular letter) and the standard references two financial-juridical criteria in Belgian law for board GC disclosures (whether retained earnings is negative and whether net income is negative for the last two years).

The GC standard derives from activities of the Belgian High Council for Auditing and Accountancy (High Council) and the Belgian Ministry of Justice. Per their 1994–1995 Annual Report, following widespread criticism of auditors, the High Council advised the Institute of Auditors to formulate a GC audit standard.[8] In response, the Institute of Auditors held a forum in the autumn of 1995, but, otherwise, did not move very quickly to propose a change in professional standards. Perhaps because of this inaction, the High Council put forth its own proposal in a written recommendation on June 20, 1996 and made certain suggestions to the Belgian Ministry of Justice, which was concurrently revising the laws on bankruptcy and creditor composition. Among the High Council's suggestions was that auditors should draw attention to statutory and regulatory provisions regarding companies in distress, and that if the auditor has substantial doubt regarding a company's continued existence, then he/she should report this doubt to the board. Following the High Council's advice, on August 8, 1997, the Ministry of Justice revised Belgian bankruptcy law and the law on creditors' composition by promulgating, as Article 77 of Belgian Company Law, two financial-juridical criteria for board of directors' GC disclosures:

> If the balance sheet shows an accumulated loss [i.e., a deficit in retained earnings], or if the profit and loss account shows a bottom-line loss in two successive years, then the Directors are required to justify in the annual statements the application of the valuation rules in the assumption of continuity.

In addition, consistent with the High Council's suggestion, the law on creditors' composition requires the auditor to inform the board of directors of GC problems.

Two years later, on June 29, 1999, the Institute of Auditors issued a draft audit standard, which became finalized on December 3, 1999 and effective on January 1, 2000.[9] The standard codifies the auditor's responsibility to notify the board of GC problems and also promulgates a role for the auditor with respect to the board's legal responsibilities regarding GC disclosures:

> The statutory auditor needs to ascertain whether a company with going-concern problems complies with all legal requirements. In particular, the auditor has to report going-concern problems encountered during his control activities to the Board of Directors ... When a company meets or threatens to meet one of the two hypotheses mentioned in Article 77 of Belgian Company Law, the auditor needs to assess whether it is desirable to remind the Board of Directors that this information should be included in the

---

[8] The Flemish Commission for Preventive Business Management, the Belgian Commission for Banking and Finance, and the financial press (e.g., "Auditing simply doesn't work ... in bankruptcies, the statutory auditor comes too late with the regularity of clockwork" [Wymeersch 1994]) all voiced criticism of the auditing profession.

[9] As the Annual Report of the Belgian Institute of Auditors for 1999 mentions, another motivation for the GC audit standard was the concurrent development of International Statement on Auditing (ISA) 570, *Going Concern*, and the view that the Belgian Institute could not afford to "lag behind" international standards.

> annual report or in the notes to the financial statements. (The Belgian GC audit standard, CA/3.12.1999)

The phrase "desirable to remind" refers to the case when a company threatens to meet the financial-juridical criteria. In the case when a company does meet the criteria and the board does not provide disclosures, the auditor *must* remind the board of its legal responsibility to provide adequate disclosures regarding the company's status as a GC. In addition, the auditor must assess whether the board's disclosures are adequate and appropriate. Even if the auditor deems the disclosures acceptable, he/she should still issue a GC opinion if he/she conclude there is a GC risk for the foreseeable future (i.e., at least 12 months from fiscal year-end). If the auditor deems the board's disclosures to be unacceptable, then he/she should issue a qualified opinion.[10] Given these changes to their legal responsibilities and professional audit standards, Belgian auditors arguably occupy a unique role in ensuring appropriate disclosure of potential company bankruptcy.

Rules Rather than Discretion

The essence of the rules rather than discretion debate is whether economic policy should be set according to pre-set rules or left to the judgment of decision-makers to adapt to current conditions. In a seminal paper, Kydland and Prescott (1977) demonstrate the critical role of credibility in policymaking. Credibility is important because absent a binding commitment to enforce rules, policymakers can resort to individual judgment. However, this option to exercise discretion is problematic because if economic agents anticipate it, then they may behave in ways that prevent policymakers from achieving their policy objectives ("time inconsistency" problem). Kydland and Prescott (1977) show that rules are preferable to discretion because rules can be made binding, and thereby avoid the time-inconsistency problem. Relevant to our context, Kydland and Prescott (1977) argue for a credible commitment to simple rules, and Belgium's standard references simple rules (financial-juridical criteria) that the law credibly codifies.

Another reason that the economics literature advocates rules for auditors is to restrain a tendency toward leniency that stems from a decision-maker's self-interest. Boot and Thakor (1993) formalize this self-interest in a model of a regulator's bank closure decision. The regulator seeks to maximize a weighted objective function of societal welfare and personal gain from attaining and/or protecting a reputation as a high-quality regulator. Because of the reputation protection component, Boot and Thakor (1993) show that the equilibrium closure policy is too lax and allows a bank to continue to operate when it should not.[11] While such reputation motives are present, auditors also have the obvious incentive to acquiesce because they receive fees from their clients. Consistent with this reasoning, many studies find that the presence of a GC opinion for a stressed, soon-to-be-bankrupt company varies inversely with company size, suggestive of auditor leniency toward their larger clients.

---

[10] We do not observe any audit opinions that are qualified for inadequate disclosure. In addition, as Table 4 shows, we note that the auditor's GC opinion decision is not mechanistic in nature (i.e., post-standard, seven of 53 bankrupt and nine of 34 non-bankrupt companies that violate both brightlines do not receive a GC opinion, whereas five of nine bankrupt and one of 25 non-bankrupt companies that violate neither brightline do receive a GC opinion).

[11] Similarly, the literature suggests precise accounting standards decrease aggressive reporting because they bolster the auditor's negotiating position and help them to stay independent. For example, Magee and Tseng (1990, 316) demonstrate conditions under which "cut-and-dried" standards, that "leave less room for disagreement among auditors," allow for greater auditor independence; per the partners in Gibbins et al.'s (2001) survey, clear-cut standards strengthen auditor positions when negotiating to prevent aggressive accounting; per Nelson et al. (2002), auditors recall being less likely to allow aggressive earnings management when there is a precise standard and the client has not engaged in transaction structuring; and Ng and Tan (2003) find that when precise guidance is absent, audit managers are likely to permit aggressive accounting, especially when oversight is weak.

Likely Effects of the Belgian GC Audit Standard

We use Matsumura et al.'s (1997) model of the auditor's GC decision as a lens through which to consider the likely effects of Belgium's GC standard. They specify the auditor's expected utility from rendering a GC opinion as Maximize $(p + \delta)\ W_B + (1 - p - \delta)(W_S - P_G)$ and from rendering a clean opinion as Maximize $p(W_B - P_S) + (1 - p)W_S$, where:

p = unperturbed posterior probability the client will terminate and δ $(0 \leq \delta \leq 1 - p)$ is a perturbation of this probability that the client will terminate due to the self-fulfilling prophecy effect of the auditor's decision to issue a GC opinion (see Vanstraelen 2003);
W_B = present value of future quasi rents to the auditor if the client terminates;
W_S = present value of future quasi rents to the auditor if the client continues $(W_S > W_B)$;
P_G = auditor's penalty for rendering an incorrect GC opinion (the cost of Type I error); and
P_S = penalty for rendering an incorrect clean opinion (the cost of a Type II error).

Matsumura et al. (1997) show the Nash Equilibrium is the auditor will issue a GC when its expected payoff exceeds that of a clean opinion (i.e., $(p + \delta)\ W_B + (1 - p - \delta)(W_S - P_G) > p(W_B - P_S) + (1 - p)W_S$). As such, the optimal decision rule is for the auditor to issue a GC when the unperturbed posterior probability the client will terminate, p, exceeds a certain threshold, p^*. After solving for p^*, the incumbent auditor issues a GC opinion if $p > (P_G + \delta(W_S - W_B - P_G))/P_G + P_S$.

Adapting Matsumura et al.'s (1997) framework to our context, following the issuance of the GC standard in Belgium, we argue that P_S should increase. This is because the auditor's penalty for rendering an incorrect clean opinion, particularly to clients that flunk the financial-juridical criteria, now involves the oversight of Belgium's Institute of Auditors and the specter of legal ramifications. Following Matsumura et al. (1997), an increase in P_S leads to a decrease in p^*, thereby making it more likely that p exceeds p^*. As such, Belgian auditors should be more likely to issue GC opinions following the effective date of the audit standard and, because GC opinions are more likely, the frequency of Type I (Type II) errors should increase (decrease).

Moreover, following Boot and Thakor (1993), the financial-juridical criteria should restrain the extent to which auditor self-interest influences GC decisions. Since auditor incentives to be lenient are rationally increasing in client size (i.e., the auditor's benefits of acquiescence are larger for larger clients), the inverse association between the presence of a GC opinion and client size should attenuate following the issuance of Belgium's GC audit standard.

III. SAMPLE AND DESCRIPTIVE STATISTICS

Sample

We collect a sample of stressed, private bankrupt companies and a matched sample of stressed, private non-bankrupt companies. We screen on stress because the Belgian GC standard specifically pertains to the audits of companies experiencing stress, and the audit opinion literature emphasizes the importance of conditioning analysis of auditor GC reporting on the presence of financial stress.[12] We focus on private companies because they constitute the overwhelming majority of the Belgian economy, and because auditor reporting may be more important for private companies than for public companies, given the poorer information environment surrounding private companies. We use a matched-sample design because there are thousands of stressed, private Belgian companies, and we must purchase audit opinions and board annual reports from the Belgian National Bank and hand-collect auditor data.[13] Because Carcello et al. (1997) demonstrate the peculiarity of GC reporting in the U.S. prior to SAS No. 59, we omit the period of transition to the Belgian GC standard. Since the transition begins in April 1997 (the laws on bankruptcy and creditors' composition) and ends in January 2000 (the effective date of the standard), we exclude companies that file bankruptcy from 1997 to 2000.[14]

Of the 499 companies subject to audit that file for bankruptcy during 1995–1996 and 2001–2002, we exclude 186 financial companies and nonfinancial companies for which pre-bankruptcy financial statements are unavailable.[15,16] Of the remaining 313, we eliminate 42 that do not exhibit an obvious sign of stress, which we define as negative net income, negative working capital, or negative stockholders' equity.[17] Of the remaining 271, we exclude 14 because of missing prior-year net income necessary to calculate one of the financial-juridical criteria, 19 due to missing board disclosure information, 14 because of missing audit information, two because their securities trade publicly, and one because it uses liquidation-basis accounting. Our bankrupt sample is 221 companies, 132 (89) prior to (after) the standard.[18]

We then match each stressed, private bankrupt company with a stressed, private, non-bankrupt company on the basis of year, industry, and, in the spirit of propensity-score-matching, probability of bankruptcy. Propensity-score-matching is a statistical technique that controls for firm-specific characteristics in the matching process (Rosenbaum and Rubin 1983; Dehejia and Wahba 2002) and, in our context, helps to reduce sample selection bias due to observable differences between stressed, bankrupt companies and stressed, non-bankrupt companies.[19]

---

[12] For example: per Louwers (1998, 145), "Hopwood et al. (1994) emphasize the importance of client financial distress for the auditor's going-concern decision because auditors issue going-concern disclosures only to clients experiencing difficulties and an apparent lack of financial stress may be due to management's manipulations. For these reasons, only financially 'stressed' companies were included in the final sample"; per Reynolds and Francis (2000, 390), "[f]or the going concern analysis, the initial sample ... is reduced to include only those companies that are potentially financially distressed. The reason for this additional screen is to evaluate auditor reporting decisions for a subset of companies for which the going concern report is a more salient decision"; per DeFond et al. (2002, 1255) "[a]s in prior research, we limit our analysis to a sample of financially distressed firms in evaluating the auditor's probability of issuing a first time going concern opinion ... This is because the going concern opinion decision is most salient among distressed firms."

[13] For example, 11,147 non-bankrupt Belgian companies report negative net income for 2001, or negative working capital for either 2001 or 2000, or negative stockholders' equity for either 2001 or 2000 (see also footnote 29).

[14] In Section IV, we examine the effects of excluding this period of transition to the Belgian GC audit standard.

[15] Our source for bankrupt companies is the Belgian National Bank. In Belgium, the objective of bankruptcy is liquidation (akin to Chapter 7 in the U.S.); whereas judicial composition is a relatively infrequent, alternative procedure, the objective of which is to rehabilitate and reorganize (akin to Chapter 11 in the U.S.).

[16] Following the bankruptcy prediction and going-concern literatures (e.g., Zmijewski 1984; Dopuch et al. 1987) we exclude financial companies (i.e., those with a NACE-BEL code starting with 65, 66, and 67), primarily because of industry-specific financial accounting and reporting practices.

[17] We do not use a cash flow proxy because prevailing Belgian GAAP does not require a Statement of Cash Flows.

[18] The decrease from 132 companies pre-standard to 89 post-standard mirrors the number of bankruptcies during these time periods, as the number of audited companies filing bankruptcy is 309 pre-standard and 190 post-standard (and our attrition rate is comparable at 57 percent pre-standard, or 177 of 309, and 53 percent post-standard, or 101 of 190).

We define a non-bankrupt company as one that does not file bankruptcy in the upcoming year. To increase our chances of a good match, we also consider companies with negative working capital or shareholders' equity in the prior year.[20] Because data collection is costly, for each stressed, bankrupt company we collect three stressed, non-bankrupt companies for the same year in the same industry. Of these, we choose the one with a probability of bankruptcy (Zmijewski 1984) closest to that of the stressed, bankrupt company. Our non-bankrupt sample consists of 221 companies, 132 (89) relating to the period prior to (after) the standard.

Descriptive Statistics

Table 1 presents descriptive statistics, partitioned by pre- and post-standard. *GCOpinion* equals 1 if the opinion expresses doubt regarding the client's ability to continue as a going concern, and 0 otherwise. *EMOpinion* equals 1 if the audit opinion references an emphasis of matter regarding the client's assumption of continuity (but without an expression of doubt), possible dissolution or applications of Belgian Company Law pertaining to the continuation of the entity, and 0 otherwise. The standard is associated with a significant increase in GC opinions. For the bankrupt sample, the mean of *GCOpinion* increases from 42.4 percent pre-standard to 71.9 percent post-standard (decrease in Type II errors from 57.6 percent to 28.1 percent). And, for the non-bankrupt sample, *GCOpinion* increases from 18.2 percent pre-standard to 33.7 percent post-standard (Type I errors).

Due to both research design and institutional details, the GC rates for the non-bankrupt sample are high in contrast to extant literature. In terms of research design, because we require our bankrupt sample to exhibit an obvious indicator of financial stress and then use a propensity-score-matching technique to identify our non-bankrupt sample, the companies we study exhibit rather severe distress and thus are more likely to have GC opinions. For example, the overall sample of U.S. public companies in Louwers (1998), which reports GC rates of 43.6 percent for bankrupt companies and 7.3 percent for non-bankrupt companies, has a mean bankruptcy probability (Hopwood et al. 1988) of 0.282. In contrast, our overall pre-standard sample, for which we report GC rates of 42.4 percent for bankrupt companies and 18.2 percent for non-bankrupt companies, has a mean bankruptcy probability (Zmijewski 1984) of 0.545. While Louwers (1998) does not permit comparison, this difference in bankruptcy probabilities likely stems from our respective non-bankrupt samples. In contrast to extant literature, we limit the extent to which less-distressed companies enter the sample and, thereby, inflate the denominator in calculating the Type I error rate (which consequently lowers the Type I error rate).

In terms of institutional details, there are several reasons why GC rates are higher in Belgium. First, as we mention, private companies comprise the large majority of the Belgian economy and, when distressed, likely have fewer financing and survival options than distressed public companies. Second, as we also mention, because of restrictions on auditor consulting and limits on career mobility, the benefits of auditor acquiescence in Belgium

---

[19] Our primary concern is if the bankrupt sample is much more distressed than the non-bankrupt sample, the impact of the new standard on auditor reporting decisions for bankrupt companies (the extent of the likely decrease in Type II errors) may be overstated relative to non-bankrupt companies (the extent of the likely increase in Type I errors).

[20] We do not consider companies with negative net income in the prior year, because doing so would constitute the financial-juridical criteria for two successive years of negative net income.

TABLE 1
Descriptive Statistics by Bankruptcy Status and Time Period

Panel A: Stressed Bankrupt Companies

Variable		Total	*PreStandard*	*PostStandard*	Tests of Differences
Observations		221	132	89	
GCOpinion	mean	0.543	0.424	0.719	t = 4.57[a]
EMOpinion	mean	0.054	0.053	0.056	t = 0.01
Pr(Bankrupt)	mean	0.638	0.641	0.634	t = −0.14
	median	0.692	0.669	0.716	z = 2.04[b]
Big6/5	mean	0.244	0.152	0.382	t = 3.81[a]
AuditLag	mean	5.615	5.788	5.360	t = −1.75[c]
	median	6.000	6.000	5.000	z = −1.41
Tenure	mean	0.118	0.061	0.202	t = 2.97[a]
Sales	mean	17.284	8.774	29.906	t = 0.99
	median	4.183	4.046	4.505	z = 0.44
GCDisclosures	mean	1.851	1.621	2.191	t = 3.05[a]
	median	2.000	2.000	2.000	z = 0.00
FJCriteria-RE	mean	0.873	0.886	0.854	t = −0.69
FJCriteria-NI	mean	0.633	0.629	0.640	t = 0.18

Panel B: Stressed Non-Bankrupt Companies

Variable		Total	*PreStandard*	*PostStandard*	Tests of Differences
Observations		221	132	89	
GCOpinion	mean	0.244	0.182	0.337	t = 2.56[b]
EMOpinion	mean	0.077	0.061	0.101	t = 1.06
Pr(Bankrupt)	mean	0.473	0.450	0.508	t = 1.27
	median	0.427	0.428	0.425	z = 1.36
Big6/5	mean	0.439	0.341	0.584	t = 3.64[a]
AuditLag	mean	4.917	4.851	5.016	t = 0.70
	median	4.933	4.900	4.933	z = 2.16[b]
Tenure	mean	0.181	0.136	0.247	t = 2.02[b]
Sales	mean	7.960	5.272	11.946	t = 1.90[c]
	median	2.812	2.513	4.374	z = 2.72[a]
GCDisclosures	mean	1.176	1.114	1.270	t = 0.96
	median	1.000	1.000	1.000	z = 0.00
FJCriteria-RE	mean	0.738	0.773	0.685	t = −1.42
FJCriteria-NI	mean	0.480	0.523	0.416	t = −1.57

[a, b, c] Significant beyond the 1 percent, 5 percent, and 10 percent levels, respectively (two-sided tests).

Test of differences in means is a two-sample t-test and test of differences in medians is a Mann-Whitney U-test.

Bankrupt observations are stressed, private Belgian companies filing bankruptcy during 1995–1996 (*PreStandard*) and 2001–2002 (*PostStandard*) with necessary financial statement, audit, and board disclosure data. For these companies, we define stress as negative net income, negative working capital, or negative stockholders' equity.

Non-bankrupt observations are stressed, private Belgian companies matched with bankrupt companies on year, industry, and *Pr(Bankrupt)* with necessary financial statement, audit, and board disclosure data. For these companies, we define stress as negative net income in the current year, negative working capital in either the current or the prior year, or negative stockholders' equity in either the current or the prior year.

(*continued on next page*)

TABLE 1 (continued)

Variable Definitions:
GCOpinion = 1 if opinion states doubt about ability to continue as a going concern, and 0 otherwise;
EMOpinion = 1 if opinion states an emphasis of matter about continuity (but without a statement of doubt), possible dissolution, or applications of Belgian Company Law pertaining to the continuation of the entity, and 0 otherwise;
Pr(Bankrupt) = bankruptcy probability (Zmijewski 1984, Table 3B, 40:800 choice-based sample)
Big6/5 = 1 if a Big 6 or 5 audit firm, and 0 otherwise;
AuditLag = number of months between financial statement date and audit opinion date;
Tenure = 1 if the length of the auditor/client relation exceeds six years (i.e., if it is more than two three-year, renewable mandates), and 0 otherwise
Sales = annual sales, in millions of 1994 Euros;
GCDisclosures = number of disclosures (ranging from 0 to 6) in the annual report of the Board of Directors of circumstances that can negatively influence the development of the company, important negative events occurring after the company's fiscal year-end, applications of Belgian Company Law pertaining to continuation of the entity, or other negative events;
FJCriteria-RE = 1 if negative retained earnings, and 0 otherwise; and
FJCriteria-NI = 1 if negative net income for the last two years, and 0 otherwise.

relate to audit fees. However, Belgian companies retain their audit firm for three-year periods (making acquiescence to retain a client less likely), and most companies are small and pay small audit fees.[21] Hence, an incorrect GC is arguably less costly to auditors in Belgium. Third, all Belgian companies with more than 50 employees have a workers' council, to which the audit firm must report and to which companies have a legal obligation to provide financial information for the purpose of wage and labor conditions negotiations (Vanstraelen and Willekens 2008). Given the existence of a workers' council, a GC could conceivably help management wrest labor concessions (e.g., headcount reductions) to facilitate company survival. Therefore, a GC opinion to a company that does not soon file for bankruptcy may be *ex ante* remedial (instead of *ex post* incorrect).

Table 1 also shows a change in Belgium's audit market from pre- to post-standard, as the *Big6/5* market share goes from 15.2 percent (34.1 percent) pre-standard to 38.2 percent (58.4 percent) post-standard for the bankrupt (non-bankrupt) sample.[22] As such, the increase in frequency of GC opinions in the post-standard period may be due to an influx of larger auditors with greater incentives to protect their investments in reputation capital (DeAngelo 1981). We examine the effect of changes in *Big6/5* market share in Section IV.

GCDisclosures is the number of negative disclosures regarding the company's going-concern status appearing in the annual report of the board of directors. To operationalize the board disclosure requirements of Belgian law that the GC audit standard references, we specify *FJCriteria-RE* and *FJCriteria-NI*, variables that equal 1 if the company has negative retained earnings for the current year and negative net income for both the current and prior year, respectively, and 0 otherwise. Of note, for the bankrupt sample (Panel A of Table 1), average *GCDisclosures* increase from 1.621 pre-standard to 2.191 post-standard, despite no change in the frequency with which these companies violate the financial-juridical criteria.

---

[21] In terms of company size, mean pre-standard *Sales* for our combined sample is 7.023 million Euros ($6 million at prevailing exchange rates). As such, our sample is smaller, by orders of magnitude, than extant literature (e.g., mean sales for Reynolds and Francis' 2000 sample of 2,439 distressed U.S. companies from 1996, for which they report a GC rate of 9.2 percent, are approximately $38 million). In terms of audit fees, though not publicly available, Willekens and Gaeremynck (2005) report mean (median) Belgian audit fees for 2001 were just 10,000 (5,000) Euros.

[22] This is consistent with the increase in Big 6/5 share in Belgium in the 1990s (Willekens and Achmadi 2003).

IV. EMPIRICAL ANALYSIS

We start by assessing the effect of the GC standard on Type I errors and Type II errors (Tables 2 and 3) and the change in the frequency of GC opinions, pre- to post-standard, by audit firm type and number of financial-juridical criteria a company violates (Table 4). We then study the characteristics associated with Type I and Type II errors (Tables 5, 6, and 7). Last, we conduct an *ex post* analysis of Type I and II errors, by pre-standard and post-standard (Table 8).

The Belgian GC Standard and the Frequency of GC Opinions

We include several covariates in our audit-opinion regressions (Tables 2, 3, 6, and 7). We specify Zmijewski's (1984) probability of bankruptcy (*Pr(Bankrupt)*) to control for incremental stress. We control for audit firm type (DeAngelo 1981), time period from financial statement date to audit opinion date (Raghunandan and Rama 1995), and audit firm tenure (Knechel and Vanstraelen 2007). Because there is no disclosure of audit fees in Belgium, we specify client revenues to proxy for the effect of auditor incentives on GC reporting. Because Vanstraelen (2002) reports that the presence of a GC opinion in Belgium is positively associated with board disclosures, we include *GCDisclosures*. Equation (1a) is our baseline ordered probit regression.[23] To examine the effects of the GC standard on Type II errors (i.e., failure to render a GC to a company that files bankruptcy the next year) and Type I errors (i.e., failure to render a clean opinion to a company that does not file bankruptcy the next year), we add *PostStandard* to Equation (1a) to specify Equation (1b), both of which we estimate separately using each sample.

$$\left.\begin{array}{r} 2 \text{ if } GCOpinion_{i,t} \\ 1 \text{ if } EMOpinion_{i,t} \\ 0 \text{ Otherwise} \end{array}\right\} |Bankrupt_{i,t+1} \begin{aligned}[t] = {} & \beta_0 + \beta_1 Pr(Bankrupt)_{i,t} + \beta_2 Big6/5_{i,t} \\ & + \beta_3 AuditLag_{i,t} + \beta_4 Tenure_{i,t} \\ & + \beta_5 Ln(Sales)_{i,t} + \beta_6 GCDisclosures_{i,t} + \varepsilon \end{aligned} \qquad (1a)$$

$$\left.\begin{array}{r} 2 \text{ if } GCOpinion_{i,t} \\ 1 \text{ if } EMOpinion_{i,t} \\ 0 \text{ Otherwise} \end{array}\right\} |Bankrupt_{i,t+1} \begin{aligned}[t] = {} & \beta_0 + \beta_1 Pr(Bankrupt)_{i,t} + \beta_2 Big6/5_{i,t} \\ & + \beta_3 AuditLag_{i,t} + \beta_4 Tenure_{i,t} \\ & + \beta_5 Ln(Sales)_{i,t} + \beta_6 GCDisclosures_{i,t} \\ & + \beta_7 PostStandard_{i,t} + \varepsilon \end{aligned} \qquad (1b)$$

$$\left.\begin{array}{r} 2 \text{ if } GCOpinion_{i,t} \\ 1 \text{ if } EMOpinion_{i,t} \\ 0 \text{ Otherwise} \end{array}\right\} |Non\text{-}Bankrupt_{i,t+1} \begin{aligned}[t] = {} & \beta_0 + \beta_1 Pr(Bankrupt)_{i,t} + \beta_2 Big6/5_{i,t} \\ & + \beta_3 AuditLag_{i,t} + \beta_4 Tenure_{i,t} \\ & + \beta_5 Ln(Sales)_{i,t} + \beta_6 GCDisclosures_{i,t} + \varepsilon \end{aligned} \qquad (1a)$$

$$\left.\begin{array}{r} 2 \text{ if } GCOpinion_{i,t} \\ 1 \text{ if } EMOpinion_{i,t} \\ 0 \text{ Otherwise} \end{array}\right\} |Non\text{-}Bankrupt_{i,t+1} \begin{aligned}[t] = {} & \beta_0 + \beta_1 Pr(Bankrupt)_{i,t} + \beta_2 Big6/5_{i,t} \\ & + \beta_3 AuditLag_{i,t} + \beta_4 Tenure_{i,t} \\ & + \beta_5 Ln(Sales)_{i,t} + \beta_6 GCDisclosures_{i,t} \\ & + \beta_7 PostStandard_{i,t} + \varepsilon \end{aligned} \qquad (1b)$$

---

[23] Because *Pr(Bankrupt)* is from Zmijewski (1984), which specifies return on assets, liabilities to assets, and the current ratio, and we specify *Ln(Sales)*, our covariates encompass four (i.e., profitability, leverage, liquidity, and size) of the five factors that extant literature (e.g., Ohlson 1980) associates with bankruptcy and/or a GC. We cannot control for the fifth factor, the association between debt status and the presence of a GC and/or bankruptcy (Chen and Church 1992; Foster et al. 1998), because Belgian companies do not disclose this information. We do not control for economic trends that may affect the GC rate because macroeconomic growth in Belgium during the two sample time periods was quite comparable—see the subsection "Temporal Concerns."

where:

GCOpinion = 1 if opinion states doubt about ability to continue as a going concern, and 0 otherwise;
EMOpinion = 1 if opinion states an emphasis of matter about continuity (but without a statement of doubt), possible dissolution, or applications of Belgian Company Law pertaining to the continuation of the entity, and 0 otherwise;
Pr(Bankrupt) = bankruptcy probability (Zmijewski 1984, Table 3B, 40:800 choice-based sample);
Big6/5 = 1 if a Big 6 or 5 audit firm, and 0 otherwise;
AuditLag = number of months between financial statement date and audit opinion date;
Tenure = 1 if the length of the auditor/client relation exceeds six years (i.e., if it is more than two three-year, renewable mandates), and 0 otherwise;
Ln(Sales) = Ln(1,000 + annual sales, in millions of 1994 Euros);
GCDisclosures = number of disclosures (ranging from 0 to 6) in the annual report of the board of directors of circumstances that can negatively influence the development of the company, important negative events occurring after the company's fiscal year-end, applications of Belgian Company Law pertaining to continuation of the entity, or other negative events;[24] and
PostStandard = 1 if after the effective date of the Belgian audit standard, and 0 otherwise.

Column two of Table 2 shows the results of estimating Equation (1b) with the bankrupt company sample (column one shows the results of estimating Equation (1a)). The coefficient for *PostStandard* is positive 0.85, significant well beyond 1 percent (two-sided test), and its inclusion increases the pseudo R^2 from 12.7 percent to 17.3 percent. The marginal effect (i.e., the partial derivative of the conditional probability of a GC with respect to the variable of interest) of *PostStandard*'s coefficient is 0.33, implying a 33 percent increase in the probability of observing a GC opinion on the financial statements of a stressed, soon-to-be bankrupt company following the adoption of the GC standard. This result provides evidence of a decrease in Type II errors.[25]

Column four of Table 2 presents the results of estimating Equation (1b) with the non-bankrupt sample (column three shows the results of estimating Equation (1a)). For this sample, *PostStandard*'s coefficient is positive 0.50, significant beyond 5 percent (two-sided test), and its inclusion increases the pseudo R^2 from 20.9 percent to 22.6 percent. *PostStandard*'s marginal effect is 0.13, implying a 13 percent increase in the probability of a

---

[24] For example: "In order to bring the number of employees in balance with the expected turnover for next year, a procedure has been started to collectively dismiss 30 employees"; and "The current crisis in the telecommunications industry will inevitably influence the future activities of the company. Restructuring measures will be necessary to deal with the economic recession. We will have to make more use of outsourcing to respond in a more flexible way to fast changing market trends." Belgian Company Law requires that if net assets are less than 50 percent of subscribed capital, the Board must decide on reorganization plans and whether to continue the entity.

[25] Per Table 5, mean and median values of *(PrBankrupt)* for the 56 bankrupt companies with GCs in the pre-standard period are 0.738 and 0.831, respectively. In contrast, mean and median values of *(PrBankrupt)* for the 64 bankrupt companies with GCs in the post-standard period are 0.722 and 0.905, respectively. As such, there is no evidence that bankrupt companies with GCs manifest lower *ex ante* distress in the post-standard period. This does not support a "self-fulfilling prophecy" effect underlying the decrease in Type II errors in the post-standard period.

TABLE 2
Audit Opinions for Stressed Bankrupt and Non-Bankrupt Companies

$$\left.\begin{array}{r} 2 \text{ if } GCOpinion_{i,t} \\ 1 \text{ if } EMOpinion_{i,t} \\ 0 \text{ Otherwise} \end{array}\right\} \begin{array}{l} |Bankrupt_{i,t+1} = \beta_0 + \beta_1 Pr(Bankrupt)_{i,t} + \beta_2 Big6/5_{i,t} \\ \qquad + \beta_3 AuditLag_{i,t} + \beta_4 Tenure_{i,t} + \beta_5 Ln(Sales)_{i,t} \\ \qquad + \beta_6 GCDisclosures_{i,t} + \varepsilon \end{array} \qquad \text{(1a)}$$

$$\left.\begin{array}{r} 2 \text{ if } GCOpinion_{i,t} \\ 1 \text{ if } EMOpinion_{i,t} \\ 0 \text{ Otherwise} \end{array}\right\} \begin{array}{l} |Bankrupt_{i,t+1} = \beta_0 + \beta_1 Pr(Bankrupt)_{i,t} + \beta_2 Big6/5_{i,t} \\ \qquad + \beta_3 AuditLag_{i,t} + \beta_4 Tenure_{i,t} + \beta_5 Ln(Sales)_{i,t} \\ \qquad + \beta_6 GCDisclosures_{i,t} + \beta_7 PostStandard_{i,t} + \varepsilon \end{array} \qquad \text{(1b)}$$

$$\left.\begin{array}{r} 2 \text{ if } GCOpinion_{i,t} \\ 1 \text{ if } EMOpinion_{i,t} \\ 0 \text{ Otherwise} \end{array}\right\} \begin{array}{l} |Non\text{-}Bankrupt_{i,t+1} = \beta_0 + \beta_1 Pr(Bankrupt)_{i,t} + \beta_2 Big6/5_{i,t} \\ \qquad + \beta_3 AuditLag_{i,t} + \beta_4 Tenure_{i,t} + \beta_5 Ln(Sales)_{i,t} \\ \qquad + \beta_6 GCDisclosures_{i,t} + \varepsilon \end{array} \qquad \text{(1a)}$$

$$\left.\begin{array}{r} 2 \text{ if } GCOpinion_{i,t} \\ 1 \text{ if } EMOpinion_{i,t} \\ 0 \text{ Otherwise} \end{array}\right\} \begin{array}{l} |Non\text{-}Bankrupt_{i,t+1} = \beta_0 + \beta_1 Pr(Bankrupt)_{i,t} + \beta_2 Big6/5_{i,t} \\ \qquad + \beta_3 AuditLag_{i,t} + \beta_4 Tenure_{i,t} + \beta_5 Ln(Sales)_{i,t} \\ \qquad + \beta_6 GCDisclosures_{i,t} + \beta_7 PostStandard_{i,t} + \varepsilon \end{array} \qquad \text{(1b)}$$

		Bankrupt		Non-Bankrupt	
Variable	**E[sign]**	**Equation (1a)**	**Equation (1b)**	**Equation (1a)**	**Equation (1b)**
Constant		0.87	0.34	−1.43	−1.27
		(0.97)	(0.36)	(−1.87)[c]	(−1.72)[c]
Pr(Bankrupt)	+	1.29	1.43	1.86	1.84
		(4.85)[a]	(5.08)[a]	(5.67)[a]	(5.50)[a]
Big6/5	+/−	0.29	0.03	0.35	0.24
		(1.24)	(0.12)	(1.75)[c]	(1.15)
AuditLag	+	−0.04	−0.01	0.02	0.00
		(−0.62)	(−0.19)	(0.25)	(0.05)
Tenure	−	0.16	−0.15	−0.12	−0.24
		(0.58)	(−0.51)	(−0.45)	(−0.80)
Ln(Sales)	−	−0.12	−0.11	−0.04	−0.06
		(−2.17)[b]	(−1.96)[b]	(−0.93)	(−1.33)
GCDisclosures	+	0.26	0.23	0.25	0.25
		(3.78)[a]	(3.08)[a]	(2.80)[a]	(2.73)[a]
PostStandard	+		0.85		0.50
			(4.10)[a]		(2.38)[b]
Observations		221	221	221	221
Pseudo R^{2d}		12.7%	17.3%	20.9%	22.6%

[a, b, c] Significant beyond the 1 percent, 5 percent, and 10 percent levels, respectively (two-sided tests).

[d] 1 − (ln*Lmle*/ln*Lconstant*), where *Lmle* is the likelihood function at the maximum-likelihood estimates and *Lconstant* is the likelihood function when all coefficients (not the constant) are constrained to equal 0.

Variable Definition:

PostStandard = 1 if after effective date of the Belgian GC audit standard, and 0 otherwise.

See Table 1 for sample information and other variable definitions. Estimation is ordered probit regression.

GC on the financial statements of a company that does not file bankruptcy in the upcoming year during the period following the GC standard. This result provides evidence of an increase in Type I errors. While *PostStandard*'s coefficient is smaller than it is for the bankrupt sample (0.50 versus 0.85), the difference is not significant, with an asymptotic t-test of −1.17 (Ben-Akiva and Lerman 1985).

The Belgian GC Standard and the Frequency of GC Opinions by Audit Firm Type

Table 3 shows the results of running essentially the same sets of regressions, but using samples we restrict to clients of Big 6/5 or non-Big 6/5 auditors. Per column four (six) of Table 3, *PostStandard*'s coefficient for the bankrupt (non-bankrupt) company sample in Table 2 largely stems from non-Big 6/5 (Big 6/5) audit firms. This suggests differentiation between auditor type and the presence of a GC around the time the audit standard was put into place. To examine this issue, Table 4 provides descriptive information on the change in GC rates by bankruptcy status, number of financial-juridical criteria violated, and auditor type. The totals of the first two columns of Table 4 agree with Table 1 (e.g., for bankrupt, the GC rate increases from 42.4 percent pre-standard to 71.9 percent post-standard; and for non-bankrupt, the GC rate increases from 18.2 percent to 33.7 percent).

Table 4 depicts three key findings. First, for both bankrupt and non-bankrupt companies during pre-standard and post-standard periods, the frequency of going-concern and emphasis-of-matter opinions increases with the number of financial-juridical criteria violated. Second, whether none, one, or both criteria are violated, these rates increase pre-standard to post-standard. Third, due to the increase in GCs by the non-Big 6/5 (from 40.2 percent pre-standard to 76.4 percent post-standard), the non-Big 6/5 are no longer more likely than the Big 6/5 to commit a Type II error following the standard. Of note, post-standard, the non-Big 6/5 often render a GC to an *ex post* bankrupt company that is in violation of both criteria (25 of 29, or 86.2 percent).

Bankrupt and Non-Bankrupt Companies—By Opinion Type by Time Period

In this section, we estimate separate GC opinion regressions for the pre- and post-standard periods. We do so to assess parameter stability across the regime change, particularly regarding our proxy for the effects of auditor incentives on their opinion decisions (*Ln(Sales)*).

For reference, Table 5 presents descriptive statistics by opinion type and time period. Of note, during the pre-standard period, *Sales* for bankrupt companies with GCs are less than *Sales* for companies with clean opinions (8.005 mean and 3.301 median million Euros versus 8.313 mean and 5.078 median million Euros, respectively); however, post-standard this is no longer the case (36.288 mean and 4.286 median million Euros versus 15.146 mean and 3.929 median million Euros, respectively). This finding is consistent with the view that a rules-based standard mitigates the influence of auditor self-interest on GC decisions.

Because the GC standard references the two financial-juridical criteria for board GC disclosures, we start by examining the relation between these disclosures and whether a company violates these financial statement thresholds, before and after the GC audit standard. Because *GCDisclosures* is a non-negative integer, we estimate Equation (2) via Poisson regression.[26]

---

[26] For bankrupt and non-bankrupt samples, the mean and variance of *GCDisclosures* are quite close (for bankrupt, the mean is 1.851 and the variance is 1.882; and for non-bankrupt, the mean is 1.176 and the variance is 1.282), thereby satisfying the "mean equals variance" (equidispersion) feature of the Poisson distribution (Winkelman 1997).

TABLE 3
Audit Opinions for Stressed Bankrupt and Non-Bankrupt Companies by Auditor Type

$$\left.\begin{matrix} 2 \text{ if } GCOpinion_{i,t} \\ 1 \text{ if } EMOpinion_{i,t} \\ 0 \text{ Otherwise} \end{matrix}\right\} |Bankrupt_{i,t+1} = \beta_0 + \beta_1 Pr(Bankrupt)_{i,t} + \beta_2 AuditLag_{i,t} + \beta_3 Tenure_{i,t} + \beta_4 Ln(Sales)_{i,t} + \beta_5 GCDisclosures_{i,t} + \varepsilon \qquad (1c)$$

$$\left.\begin{matrix} 2 \text{ if } GCOpinion_{i,t} \\ 1 \text{ if } EMOpinion_{i,t} \\ 0 \text{ Otherwise} \end{matrix}\right\} |Bankrupt_{i,t+1} = \beta_0 + \beta_1 Pr(Bankrupt)_{i,t} + \beta_2 AuditLag_{i,t} + \beta_3 Tenure_{i,t} + \beta_4 Ln(Sales)_{i,t} + \beta_5 GCDisclosures_{i,t} + \beta_6 PostStandard_{i,t} + \varepsilon \qquad (1d)$$

$$\left.\begin{matrix} 2 \text{ if } GCOpinion_{i,t} \\ 1 \text{ if } EMOpinion_{i,t} \\ 0 \text{ Otherwise} \end{matrix}\right\} |Non\text{-}Bankrupt_{i,t+1} = \beta_0 + \beta_1 Pr(Bankrupt)_{i,t} + \beta_2 AuditLag_{i,t} + \beta_3 Tenure_{i,t} + \beta_4 Ln(Sales)_{i,t} + \beta_5 GCDisclosures_{i,t} + \varepsilon \qquad (1c)$$

$$\left.\begin{matrix} 2 \text{ if } GCOpinion_{i,t} \\ 1 \text{ if } EMOpinion_{i,t} \\ 0 \text{ Otherwise} \end{matrix}\right\} |Non\text{-}Bankrupt_{i,t+1} = \beta_0 + \beta_1 Pr(Bankrupt)_{i,t} + \beta_2 AuditLag_{i,t} + \beta_3 Tenure_{i,t} + \beta_4 Ln(Sales)_{i,t} + \beta_5 GCDisclosures_{i,t} + \beta_6 PostStandard_{i,t} + \varepsilon \qquad (1d)$$

		Stressed Bankrupt				Stressed Non-Bankrupt			
		Big 6/5		Non-Big 6/5		Big 6/5		Non-Big 6/5	
Variable	E[sign]	Equation (1c)	Equation (1d)	Equation (1c)	Equation (1d)	Equation (1c)	Equation (1d)	Equation (1c)	Equation (1d)
Constant		1.23	1.34	0.70	−0.19	−0.97	−1.25	−1.44	−1.12
		(0.63)	(0.68)	(0.66)	(−0.17)	(−0.81)	(−1.06)	(−1.38)	(−1.06)
Pr(Bankrupt)	+	1.85	1.90	1.22	1.55	1.61	1.76	2.23	2.14
		(2.91)[a]	(2.86)[a]	(3.96)[a]	(4.53)[a]	(3.23)[a]	(3.17)[a]	(4.34)[a]	(4.11)[a]

(*continued on next page*)

TABLE 3 (continued)

		Stressed Bankrupt				Stressed Non-Bankrupt			
		Big 6/5		Non-Big 6/5		Big 6/5		Non-Big 6/5	
Variable	E[sign]	Equation (1c)	Equation (1d)	Equation (1c)	Equation (1d)	Equation (1c)	Equation (1d)	Equation (1c)	Equation (1d)
AuditLag	+	−0.11	−0.11	0.02	0.08	−0.02	−0.05	0.08	0.08
		(−1.17)	(−1.20)	(0.27)	(1.01)	(−0.19)	(−0.46)	(0.79)	(0.75)
Tenure	−	0.26	0.32	0.15	−0.26	0.23	0.15	−0.38	−0.49
		(0.09)	(0.11)	(0.54)	(−0.83)	(0.63)	(0.37)	(−0.78)	(−0.94)
Ln(Sales)	−	−0.15	−0.15	−0.12	−0.11	−0.04	−0.04	−0.06	−0.09
		(−1.33)	(−1.36)	(−1.83)^{c}	(−1.73)^{c}	(−0.59)	(−0.57)	(−0.95)	(−1.27)
GCDisclosures	+	0.56	0.58	0.20	0.15	0.37	0.35	0.16	0.17
		(3.00)^{a}	(2.94)^{a}	(2.61)^{b}	(1.76)^{c}	(2.80)^{a}	(2.61)^{a}	(1.21)	(1.21)
PostStandard	+		−0.15		1.24		0.60		0.44
			(−0.34)		(4.83)^{a}		(2.09)^{b}		(1.31)
Observations		54	54	167	167	97	97	124	124
Pseudo R^{2d}		25.7%	25.9%	10.0%	18.9%	17.3%	20.0%	21.7%	22.8%

[a,b,c] Significant beyond the 1 percent, 5 percent, and 10 percent levels, respectively (two-sided tests).

[d] 1 − (ln*Lmle*/ln*Lconstant*), where *Lmle* is the likelihood function at the maximum-likelihood estimates and *Lconstant* is the likelihood function when all coefficients (not the constant) are constrained to equal 0.

Variable Definition:

PostStandard = 1 if after effective date of the Belgian GC audit standard, and 0 otherwise.

See Table 1 for sample information and other variable definitions. Estimation is ordered probit regression.

TABLE 4

Audit Opinion Rates by Bankruptcy Status, Auditor Type, Time Period, and Number of Financial-Juridical Criteria Violated

Panel A: Stressed Bankrupt Companies

	Total		Violate Neither FJ Criteria		Violate One FJ Criteria		Violate Both FJ Criteria	
	PreStandard	*PostStandard*	*PreStandard*	*PostStandard*	*PreStandard*	*PostStandard*	*PreStandard*	*PostStandard*
Big 6/5								
GCOpinion	55.0% (11/20)	64.7% (22/34)	0.0% (0/3)	NA (0/0)	57.1% (4/7)	50.0% (5/10)	70.0% (7/10)	70.8% (17/24)
EMOpinion	5.0% (1/20)	8.8% (3/34)	0.0% (0/3)	NA (0/0)	0.0% (0/7)	10.0% (1/10)	10.0% (1/10)	8.3% (2/24)
Non-Big 6/5								
GCOpinion	40.2% (45/112)	76.4% (42/55)	10.0% (1/10)	55.6% (5/9)	29.0% (9/31)	70.6% (12/17)	49.3% (35/71)	86.2% (25/29)
EMOpinion	5.4% (6/112)	3.6% (2/55)	0.0% (0/10)	0.0% (0/9)	6.5% (2/31)	0.0% (0/17)	5.6% (4/71)	6.9% (2/29)
Total								
GCOpinion	42.4% (56/132)	71.9% (64/89)	7.7% (1/13)	55.6% (5/9)	34.2% (13/38)	63.0% (17/27)	51.9% (42/81)	79.2% (42/53)
EMOpinion	5.3% (7/132)	5.6% (5/89)	0.0% (0/13)	0.0% (0/9)	5.3% (2/38)	5.9% (1/27)	6.2% (5/81)	7.5% (4/53)

Panel B: Stressed Non-Bankrupt Companies

	Total		Violate Neither FJ Criteria		Violate One FJ Criteria		Violate Both FJ Criteria	
	PreStandard	*PostStandard*	*PreStandard*	*PostStandard*	*PreStandard*	*PostStandard*	*PreStandard*	*PostStandard*
Big 6/5								
GCOpinion	26.7% (12/45)	42.3% (22/52)	0.0% (0/6)	0.0% (0/11)	42.9% (6/14)	41.2% (7/17)	24.0% (6/25)	62.5% (15/24)
EMOpinion	6.7% (3/45)	11.5% (6/52)	0.0% (0/6)	9.1% (1/11)	7.1% (1/14)	5.9% (1/17)	8.0% (2/25)	16.7% (4/24)
Non-Big 6/5								
GCOpinion	13.8% (12/87)	21.6% (8/37)	0.0% (0/17)	0.0% (0/14)	15.2% (5/33)	23.1% (3/13)	18.9% (7/37)	50.0% (5/10)
EMOpinion	5.7% (5/87)	8.1% (3/37)	0.0% (0/17)	0.0% (0/14)	6.1% (2/33)	15.4% (2/13)	8.1% (3/37)	10.0% (1/10)
Total								
GCOpinion	18.2% (24/132)	33.7% (30/89)	0.0% (0/23)	0.0% (0/25)	23.4% (11/47)	33.3% (10/30)	21.0% (13/62)	58.8% (20/34)
EMOpinion	6.1% (8/132)	10.1% (9/89)	0.0% (0/23)	4.0% (1/25)	6.4% (3/47)	10.0% (3/30)	8.1% (5/62)	14.7% (5/34)

FJ Criteria pertains to the financial juridical criteria that Belgian Company Law specifies and Belgian audit standard references (i.e., whether retained earnings are negative for the current year and whether net income is negative for the current and prior years).

See Table 1 for sample information and variable definitions.

TABLE 5
Characteristics of Type I and II Audit Opinion Errors by Time Period

			PreStandard		*PostStandard*	
			$\mathit{Bankrupt}_{t+1}$		$\mathit{Bankrupt}_{t+1}$	
			No	Yes	No	Yes
CleanOpinion	Observations		100	69 (Type II)	50	20 (Type II)
	$\mathit{Pr(Bankrupt)}_t$	mean	0.359	0.529	0.346	0.367
		median	0.322	0.555	0.305	0.284
	$\mathit{Big6/5}_t$	mean	0.300	0.116	0.480	0.450
	$\mathit{AuditLag}_t$	mean	4.808	5.696	4.872	5.700
		median	4.900	6.000	4.700	5.000
	Tenure_t	mean	0.150	0.087	0.280	0.250
	Sales_t	mean	6.193	8.313	13.137	15.146
		median	3.127	5.078	6.580	3.929
	$\mathit{GCDisclosures}_t$	mean	1.000	1.420	0.840	1.400
		median	1.000	1.000	1.000	1.000
	$\mathit{FJCriteria\text{-}RE}_t$	mean	0.710	0.797	0.460	0.700
	$\mathit{FJCriteria\text{-}NI}_t$	mean	0.500	0.522	0.240	0.450
EMOpinion	Observations		8	7	9	5
	$\mathit{Pr(Bankrupt)}_t$	mean	0.732	0.968	0.713	0.578
		median	0.996	0.994	0.670	0.671
	$\mathit{Big6/5}_t$	mean	0.375	0.143	0.667	0.600
	$\mathit{AuditLag}_t$	mean	4.113	5.571	5.944	5.200
		median	4.500	6.000	5.000	5.000
	Tenure_t	mean	0.250	0.000	0.222	0.000
	Sales_t	mean	0.369	19.463	2.942	7.252
		median	0.083	2.315	2.656	8.589
	$\mathit{GCDisclosures}_t$	mean	1.250	1.429	1.889	1.800
		median	1.000	1.000	1.000	2.000
	$\mathit{FJCriteria\text{-}RE}_t$	mean	1.000	1.000	0.889	0.800
	$\mathit{FJCriteria\text{-}NI}_t$	mean	0.625	0.714	0.556	1.000
GCOpinion	Observations		24 (Type I)	56	30 (Type I)	64
	$\mathit{Pr(Bankrupt)}_t$	mean	0.733	0.738	0.716	0.722
		median	0.740	0.831	0.841	0.905
	$\mathit{Big6/5}_t$	mean	0.500	0.196	0.733	0.344
	$\mathit{AuditLag}_t$	mean	5.272	5.929	4.978	5.266
		median	5.000	6.000	5.000	5.000
	Tenure_t	mean	0.042	0.036	0.200	0.203
	Sales_t	mean	3.068	8.005	12.662	36.288
		median	1.112	3.301	2.406	4.286
	$\mathit{GCDisclosures}_t$	mean	1.542	1.893	1.800	2.469
		median	1.500	2.000	1.000	2.000
	$\mathit{FJCriteria\text{-}RE}_t$	mean	0.958	0.982	1.000	0.906
	$\mathit{FJCriteria\text{-}NI}_t$	mean	0.583	0.750	0.667	0.672

See Table 1 for sample information and variable definitions.

$$GCDisclosures_{it}|Bankrupt_{i,t+1} = \beta_0 + \beta_1 FJCriteria\text{-}RE_{i,t} + \beta_2 FJCriteria\text{-}NI_{i,t} + \varepsilon \tag{2}$$

where:

FJCriteria-RE = 1 if retained earnings are negative, and 0 otherwise; and
FJCriteria-NI = 1 if net income is negative for the last two years, and 0 otherwise.

The results in columns one and five of Table 6 show that in both periods the extent of disclosures is positively associated with the negative retained earnings criteria. This finding suggests that the change to Belgian law had little direct effect on board disclosures for these bankrupt companies and that the increase in *GCDisclosures* (Table 1) relates to an across-the-board intercept shift. Columns two and six of Table 6 show the results of estimating Equation (1a) for the bankrupt sample in the pre-standard and post-standard periods. Of note, *Ln(Sales)*'s coefficient is negative in the pre-standard period, confirming the size effect that extant literature documents, but is insignificant in the post-standard period. This finding suggests that a rules-based standard mitigates the extent to which self-interest influences an auditor's GC decisions.

As in columns two and six of Table 6 (and columns one and two of Table 2), the coefficients on *GCDisclosures* are positive. Vanstraelen (2002) suggests such disclosures reduce auditor-management conflict and increase the likelihood the auditor renders a GC, which implies *GCOpinion* and *GCDisclosures* are endogenous. To address potential endogeneity, we conduct Hausman's (1978) test. We first create an instrumental variable by regressing *GCDisclosures* on all exogenous variables in the system. We then add the fitted values, *GCDisclosures-IV*, to Equation (1a) to specify Equation (1e). If *GCDisclosures-IV*'s coefficient is significant in the Equation (1e) estimation, then we should then estimate Equation (1f).

$$\left.\begin{array}{r} 2 \text{ if } GCOpinion_{i,t} \\ 1 \text{ if } EMOpinion_{i,t} \\ 0 \text{ Otherwise} \end{array}\right\} |Bankrupt_{i,t+1} = \begin{array}{l} \beta_0 + \beta_1 Pr(Bankrupt)_{i,t} + \beta_2 Big6/5_{i,t} \\ + \beta_3 AuditLag_{i,t} + \beta_4 Tenure_{i,t} + \beta_5 Ln(Sales)_{i,t} \\ + \beta_6 GCDisclosures_{i,t} + \beta_7 GCDisclosures\text{-}IV_{i,t} + \varepsilon \end{array} \tag{1e}$$

$$\left.\begin{array}{r} 2 \text{ if } GCOpinion_{i,t} \\ 1 \text{ if } EMOpinion_{i,t} \\ 0 \text{ Otherwise} \end{array}\right\} |Bankrupt_{i,t+1} = \begin{array}{l} \beta_0 + \beta_1 Pr(Bankrupt)_{i,t} + \beta_2 Big6/5_{i,t} \\ + \beta_3 AuditLag_{i,t} + \beta_4 Tenure_{i,t} + \beta_5 Ln(Sales)_{i,t} \\ + \beta_6 GCDisclosures\text{-}IV_{i,t} + \varepsilon \end{array} \tag{1f}$$

where:

GCDisclosures-IV = fitted values from regressing *GCDisclosures* on all exogenous variables in the system (i.e., all right-hand-side variables in Equations (1a) and (2)), via Poisson regression.

Columns three and seven of Table 6 present the results of estimating Equation (1e) for the bankrupt sample, depending on whether their bankruptcy filing is before or after the GC audit standard. The coefficient for *GCDisclosures-IV* is positive in the pre-standard period, and, thus, we reject the hypothesis that board disclosures and audit opinions are exogenous. However, in the post-standard period, during which the increase in *GCDisclosures* appears to manifest as an across-the-board intercept shift, we cannot reject this null.

TABLE 6
Board Disclosures and Audit Opinions for Stressed Bankrupt Companies by Time Period

$$GCDisclosures_{it}|Bankrupt_{i,t+1} = \beta_0 + \beta_1 FJCriteria\text{-}RE_{i,t} + \beta_2 FJCriteria\text{-}NI_{i,t} + \varepsilon \qquad (2)$$

$$\left.\begin{array}{r}2 \text{ if } GCOpinion_{i,t}\\ 1 \text{ if } EMOpinion_{i,t}\\ 0 \text{ Otherwise}\end{array}\right\} |Bankrupt_{i,t+1} = \beta_0 + \beta_1 Pr(Bankrupt)_{i,t} + \beta_2 Big6/5_{i,t} + \beta_3 AuditLag_{i,t} + \beta_4 Tenure_{i,t} + \beta_5 Ln(Sales)_{i,t} + \beta_6 GCDisclosures_{i,t} + \varepsilon \qquad (1a)$$

$$\left.\begin{array}{r}2 \text{ if } GCOpinion_{i,t}\\ 1 \text{ if } EMOpinion_{i,t}\\ 0 \text{ Otherwise}\end{array}\right\} |Bankrupt_{i,t+1} = \beta_0 + \beta_1 Pr(Bankrupt)_{i,t} + \beta_2 Big6/5_{i,t} + \beta_3 AuditLag_{i,t} + \beta_4 Tenure_{i,t} + \beta_5 Ln(Sales)_{i,t} + \beta_6 GCDisclosures_{i,t} + \beta_7 GCDisclosures\text{-}IV_{i,t} + \varepsilon \qquad (1e)$$

$$\left.\begin{array}{r}2 \text{ if } GCOpinion_{i,t}\\ 1 \text{ if } EMOpinion_{i,t}\\ 0 \text{ Otherwise}\end{array}\right\} |Bankrupt_{i,t+1} = \beta_0 + \beta_1 Pr(Bankrupt)_{i,t} + \beta_2 Big6/5_{i,t} + \beta_3 AuditLag_{i,t} + \beta_4 Tenure_{i,t} + \beta_5 Ln(Sales)_{i,t} + \beta_6 GCDisclosures\text{-}IV_{i,t} + \varepsilon \qquad (1f)$$

		PreStandard				*PostStandard*			
Variable	**E[sign]**	**Equation (2)**	**Equation (1a)**	**Equation (1e)**	**Equation (1f)**	**Equation (2)**	**Equation (1a)**	**Equation (1e)**	**Equation (1f)**
Constant		−0.53 (−1.60)	0.64 (0.54)	0.60 (0.50)	0.47 (0.39)	0.12 (0.46)	0.60 (0.40)	0.45 (0.28)	0.48 (0.32)
FJCriteria-RE	+	0.98 (2.79)[a]				0.69 (2.51)[b]			
FJCriteria-NI	+	0.16 (1.03)				0.07 (0.45)			

(continued on next page)

TABLE 6 (continued)

Variable	E[sign]	*PreStandard*				*PostStandard*			
		Equation (2)	Equation (1a)	Equation (1e)	Equation (1f)	Equation (2)	Equation (1a)	Equation (1e)	Equation (1f)
Pr(Bankrupt)	+		1.52	1.44	1.41		1.54	1.26	1.21
			(4.12)^a	(3.84)^a	(3.72)^a		(2.80)^a	(1.74)^c	(1.75)^c
Big6/5	+		0.54	0.37	0.31		−0.49	−0.53	−0.47
			(1.71)^c	(0.98)	(0.87)		(−1.26)	(−1.26)	(−1.23)
AuditLag	+		0.04	−0.00	0.00		−0.04	0.00	−0.00
			(0.49)	(−0.01)	(0.05)		(−0.40)	(0.01)	(−0.00)
Tenure	−		−0.39	−0.40	−0.39		−0.11	−0.07	−0.09
			(−0.78)	(−0.78)	(−0.79)		(−0.28)	(−0.19)	(−0.21)
Ln(Sales)	−		−0.15	−0.20	−0.19		−0.05	−0.08	−0.08
			(−2.19)^b	(−2.90)^a	(−2.71)^a		(−0.60)	(−0.82)	(−0.86)
GCDisclosures	+		0.20	0.14			0.30	0.27	
			(2.10)^b	(1.44)			(2.17)^b	(1.68)^c	
GCDisclosures-IV	+			0.70	0.83			0.27	0.50
				(2.16)^b	(2.63)^a			(0.54)	(1.26)
Observations		132	132	132	132	89	89	89	89
Pseudo R^{2d}		NA	12.4%	14.3%	13.3%	NA	18.8%	19.2%	16.1%

[a,b,c] Significant beyond the 1 percent, 5 percent, and 10 percent levels, respectively (two-sided tests).

[d] 1 − (ln*Lmle*/ln*Lconstant*), where *Lmle* is the likelihood function at the maximum-likelihood estimates and *Lconstant* is the likelihood function when all coefficients (not the constant) are constrained to equal 0.

Variable Definition:

GCDisclosures-IV = fitted values from regressing *GCDisclosures* on all exogenous variables in the system (i.e., all right-hand-side variables in Equations (1a) and (2)).

See Table 1 for sample information and other variable definitions. Estimation is Poisson regression for Equation (2) and ordered probit regression for Equations (1a), (1e), and (1f).

Importantly, though, and regardless of specification, *Ln(Sales)*'s coefficient is negative pre-standard, but insignificant post-standard.[27]

We also estimate Equations (2), (1a), (1e), and (1f) using the non-bankrupt sample. As columns one and five of Table 7 show, for these companies, *GCDisclosures* is positively associated with whether a company has negative retained earnings, but is significant only in the post-standard period. Prior to the change in Belgian law, there is no discernible relation between the extent of board disclosures and the financial-juridical criteria. Subsequent to this change in law, an association between disclosures and whether a company has negative retained earnings emerges. Therefore, for non-bankrupt companies, the law appears to have led to a relation between board disclosures and the retained earnings criteria. However, because these companies did not file bankruptcy the next year, this relation did not provide an early warning of failure.

As columns two and six of Table 7 show, there is also some evidence of a size effect for the non-bankrupt companies in the pre-standard period, but no such evidence post-standard. Again, these findings support the view that a shift toward a rules-based GC standard mitigates the extent to which auditor self-interest influences the GC opinion decision.[28] Also of note, there is evidence in the post-standard period that board GC disclosures and auditor GC opinions are endogenous. This result suggests that the increase in Type I errors following the effective date of the standard may stem, perhaps, from auditor over-reliance on (an increase in) board disclosures.

Temporal Concerns

As we state in Section III, we exclude companies if they file for bankruptcy during the period of transition to the GC standard. Our rationale for excluding these companies is that during the period of transition from SAS No. 34 to SAS No. 59 in the U.S., the rate of GCs for *ex post* bankrupt companies was peculiarly low (Carcello et al. 1997). However, the transition from discretionary circular letter to rules-based audit standard in Belgium extends considerably longer than the 21 months preceding SAS No. 59 in the U.S. In Belgium, the transition period to the standard begins in April 1997 with the introduction of a new bankruptcy law and ends with the effective date of the GC standard on January 1, 2000. Because the GC standard references criteria that Belgian Company Law and law on creditors' composition codify, it is important to assess GC reporting during Belgium's transition period because the increase in GCs may stem from changes in Belgian law, and not from the Belgian GC audit standard.

To examine this issue, we identify the 202 stressed, private Belgian companies filing for bankruptcy during the period January 1, 1997 to December 31, 2000 that do not use liquidation-basis accounting. Overall, 54.4 percent (110 of 202) of these companies have a GC opinion on their pre-bankruptcy financial statements. Consistent with an effect stemming from the changes to Belgian law, the GC rate increases somewhat throughout the transition period; from 49.2 percent to 58.1 percent to 56.0 percent to 56.9 percent for companies filing bankruptcy in 1997, 1998, 1999, and 2000, respectively. While the transition rate of 54.4 percent is higher than the pre-standard rate of 42.4 percent, it is lower

---

[27] Among the bankrupt companies, Sabena Airlines is (by far) the largest, is quite stressed (*Pr(Bankrupt)* of 0.893), and received a GC from its Big 6/5 audit firm in the post-standard period. Our results do not change if we remove this observation and re-estimate the Table 2 and 6 regressions.

[28] Among the non-bankrupt companies, Societe Nationale Des Chemins de Fer is (by far) the largest, is not very stressed (*Pr(Bankrupt)* of 0.053) and received a clean opinion from its non-Big 6/5 audit firm in the post-standard period. Our results do not change if we remove this observation and re-estimate the Table 2 and Table 7 regressions.

TABLE 7
Board Disclosures and Audit Opinions for Stressed Non-Bankrupt Companies by Time Period

$$GCDisclosures_{it}|Non\text{-}Bankrupt_{i,t+1} = \beta_0 + \beta_1 FJCriteria\text{-}RE_{i,t} + \beta_2 FJCriteria\text{-}NI_{i,t} + \varepsilon \qquad (2)$$

$$\left.\begin{array}{l}2 \text{ if } GCOpinion_{i,t}\\ 1 \text{ if } EMOpinion_{i,t}\\ \quad 0 \text{ Otherwise}\end{array}\right\} |Non\text{-}Bankrupt_{i,t+1} = \beta_0 + \beta_1 Pr(Bankrupt)_{i,t} + \beta_2 Big6/5_{i,t} + \beta_3 AuditLag_{i,t} + \beta_4 Tenure_{i,t} + \beta_5 Ln(Sales)_{i,t} + \beta_6 GCDisclosures_{i,t} + \varepsilon \qquad (1a)$$

$$\left.\begin{array}{l}2 \text{ if } GCOpinion_{i,t}\\ 1 \text{ if } EMOpinion_{i,t}\\ \quad 0 \text{ Otherwise}\end{array}\right\} |Non\text{-}Bankrupt_{i,t+1} = \beta_0 + \beta_1 Pr(Bankrupt)_{i,t} + \beta_2 Big6/5_{i,t} + \beta_3 AuditLag_{i,t} + \beta_4 Tenure_{i,t} + \beta_5 Ln(Sales)_{i,t} + \beta_6 GCDisclosures_{i,t} + \beta_7 GCDisclosures\text{-}IV_{i,t} + \varepsilon \qquad (1e)$$

$$\left.\begin{array}{l}2 \text{ if } GCOpinion_{i,t}\\ 1 \text{ if } EMOpinion_{i,t}\\ \quad 0 \text{ Otherwise}\end{array}\right\} |Non\text{-}Bankrupt_{i,t+1} = \beta_0 + \beta_1 Pr(Bankrupt)_{i,t} + \beta_2 Big6/5_{i,t} + \beta_3 AuditLag_{i,t} + \beta_4 Tenure_{i,t} + \beta_5 Ln(Sales)_{i,t} + \beta_6 GCDisclosures\text{-}IV_{i,t} + \varepsilon \qquad (1f)$$

		PreStandard				*PostStandard*			
Variable	**E[sign]**	**Equation (2)**	**Equation (1a)**	**Equation (1e)**	**Equation (1f)**	**Equation (2)**	**Equation (1a)**	**Equation (1e)**	**Equation (1f)**
Constant		−0.28	−0.68	−0.69	−0.61	−0.76	−2.42	−3.28	−3.36
		(−1.34)	(−0.65)	(−0.65)	(−0.60)	(−2.74)^a	(−1.77)^c	(−2.03)^b	(−2.05)^b
FJCriteria-RE	+	0.32				1.28			
		(1.38)				(4.17)^a			
FJCriteria-NI	+	0.23				−0.04			
		(1.32)				(−0.22)			

(continued on next page)

TABLE 7 (continued)

		PreStandard				*PostStandard*			
Variable	**E[sign]**	**Equation (2)**	**Equation (1a)**	**Equation (1e)**	**Equation (1f)**	**Equation (2)**	**Equation (1a)**	**Equation (1e)**	**Equation (1f)**
Pr(Bankrupt)	+		1.96	2.84	2.91		2.29	1.90	1.96
			(3.69)[a]	(2.27)[b]	(2.40)[b]		(3.85)[a]	(2.86)[a]	(2.89)[a]
Big6/5	−		0.16	0.12	0.12		0.56	0.28	0.30
			(0.51)	(0.39)	(0.38)		(1.76)[c]	(0.79)	(0.83)
AuditLag	+		0.15	0.08	0.07		−0.09	−0.12	−0.12
			(1.82)[c]	(0.57)	(0.54)		(−0.79)	(−1.10)	(−1.15)
Tenure	−		−0.80	−0.83	−0.81		0.11	−0.46	−0.46
			(−1.21)	(−1.22)	(−1.23)		(0.31)	(−0.77)	(−0.78)
Ln(Sales)	−		−0.15	−0.08	−0.07		0.05	0.03	0.04
			(−2.00)[b]	(−0.58)	(−0.58)		(0.73)	(0.43)	(0.48)
GCDisclosures	+		0.21	0.21			0.29	0.14	
			(1.29)	(1.33)			(2.38)[b]	(1.03)	
GCDisclosures-IV	+			−0.89	−0.79			1.40	1.53
				(−0.76)	(−0.70)			(2.28)[b]	(2.61)[a]
Observations		132	132	132	132	89	89	89	89
Pseudo R^2[d]		NA	24.3%	24.8%	23.7%	NA	23.6%	31.2%	30.5%

[a,b,c] Significant beyond the 1 percent, 5 percent, and 10 percent levels, respectively (two-sided tests).

[d] 1 − (ln*Lmle*/n*Lconstant*), where *Lmle* is the likelihood function at the maximum-likelihood estimates and *Lconstant* is the likelihood function when all coefficients (not the constant) are constrained to equal 0.

Variable Definition:

GCDisclosures-IV = fitted values from regressing *GCDisclosures* on all exogenous variables in the system (i.e., all right-hand-side variables in Equations (1a) and (2)).

See Table 1 for sample information and other variable definitions. Estimation is Poisson regression for equation (2) and ordered probit regression for Equations (1a), (1e). and (1f).

than the post-standard rate of 71.9 percent (Table 1, panel A). A z-test of the change in proportions rejects a one-tailed null of no difference between the transition and pre-standard periods beyond 5 percent, and between the transition and post-standard periods beyond 0.1 percent. As such, the increase in GCs for bankrupt companies is considerably stronger during the post-standard period.

It is also possible that the increase in GC opinions, from pre-standard to post-standard, is (at least partly) due to the choice of these two time periods. Even if firm fundamentals do not change, auditors seem likely to render more GC opinions in a recession than in an expansion. However, from the standpoint of the Belgian economy, the two time periods are quite similar (e.g., average growth in per capita GDP for 1995–1996 and 2000–2001 was 2.65 percent and 2.45 percent, respectively). Moreover, as we document in footnote 18, Dewaelheyns and Van Hulle (2008) report a *decrease* in Belgian bankruptcy rates subsequent to the 1997 bankruptcy code reform; a finding that stands directly at odds with the increase in GC audit opinions we document.

Consequences of the Increase in Type I Errors and the Decrease in Type II Errors

Overall, our findings indicate that the GC standard results in an increase in Type I errors and a decrease in Type II errors. While we acknowledge that whether the costs of more Type I errors exceed the benefits of less Type II errors is a matter of debate,[29] Table 8 attempts an *ex post* analysis of these audit opinion errors, by pre-standard versus post-standard periods.

For creditors, a GC can serve as an early, or a false, warning of impending bankruptcy. Whereas in the pre-standard period, the GC was an early warning for 42 percent (56 of 132) of sample bankruptcies, representing 34 percent (311.920/(311.920 + 169.498 + 430.215)) of liabilities; post-standard, it warns in 72 percent (64 of 89) of bankruptcies, representing 92 percent (1,788.608/(1,788.608 + 13.175 + 144.880)) of liabilities. Offsetting the benefits of this enhancement in early warning, however, is that the degree to which GC opinions falsely warn creditors also increases. Whereas in the pre-standard period, the opinion was a false warning for 18 percent (24 of 132) of our non-bankrupt sample, representing 12 percent (51.216/(51.216 + 7.160 + 378.000)) of total liabilities; post-standard, this false warning increases to 34 percent (30 of 89) and represents 34 percent (212.850/(212.850 + 43.416 + 374.500)) of total liabilities, respectively. Overall, from a creditor standpoint, the GC standard seems beneficial.

For auditors, reputation effects aside, the primary cost of a Type I error is possible client loss. Whereas pre-standard 12.5 percent (3 of 24) of companies that receive a GC but do not file bankruptcy the following year change their audit firm at the next opportunity, post-standard this rate of auditor change increases to 30 percent (9 of 30).[30] From an auditor standpoint, the standard is associated with an increase in auditor switching, likely stemming from an increase in the frequency of Type I errors.

For companies, the costs of an increase in the auditor's Type I error include a decrease in the availability, and/or an increase in the cost, of debt financing. For the pre-standard period, we are able to obtain financial statement information to compute changes in debt

---

[29] Whether the costs of more Type I errors exceed the benefits of less Type II errors is debatable especially given that the ratio of stressed bankrupt to stressed non-bankrupt companies is probably less than the 33:13 ratio of marginal effects we discuss earlier in Section IV. For example, of the 95 Belgian companies that file bankruptcy in 2002, 84 meet our bankrupt-sample definition of stress, whereas there are 10,011 non-bankrupt companies, of the 11,147 we mention in footnote 13, that meet this same definition of stress.

[30] Because Belgian companies retain audit firms for renewable, three-year "mandate" periods, we identify whether a company changes its audit firm when its existing mandate period expires (i.e., at its next opportunity).

TABLE 8
Ex Post Analysis of Type I and II Audit Opinion Errors by Time Period

			PreStandard		*PostStandard*	
			Bankrupt$_{t+1}$		*Bankrupt$_{t+1}$*	
			No	Yes	No	Yes
CleanOpinion	Observations		100	69 (Type II)	50	20 (Type II)
	Liabilities$_t$	mean	3.780	6.235	7.490	7.244
		median	2.016	4.647	3.395	4.402
		total	378.000	430.215	374.500	144.880
	Observations		67 (of 100)	61 (of 69)	43 (of 50)	13 (of 20)
	ΔDebt/Assets$_t$	mean	0.051	NA	0.800	NA
		median	−0.006	NA	−0.006	NA
	CostofDebt$_t$	mean	0.108	0.122	0.062	0.091
		median	0.093	0.113	0.059	0.068
	CostofDebt$_{t+1}$	mean	0.092	NA	0.091	NA
		median	0.089	NA	0.057	NA
	Employees$_{t-1,t}$	total	4,207 to 4,097	3,889 to 3,635	2,001 to 2,223	1,114 to 1,212
		% neg.	52% (35/67)	61% (38/61)	26% (11/43)	36% (4/13)
	Employees$_{t,t+1}$	total	4,097 to 3,937	NA	2,223 to 2,368	NA
		% neg.	57% (38/67)	NA	37% (16/43)	NA
EMOpinion	Observations		8	7	9	5
	Liabilities$_t$	mean	0.895	24.214	4.824	2.635
		median	0.148	10.323	1.990	3.393
		total	7.160	169.498	43.416	13.175
	Auditor$\Delta_{t,t+\tau}$	percent	13%	NA	11%	NA
	Observations		3 (of 8)	4 (of 7)	6 (of 9)	3 (of 5)
	ΔDebt/Assets$_t$	mean	−0.057	NA	0.173	NA
		median	0.000	NA	0.072	NA
	CostofDebt$_t$	mean	0.103	0.196	0.049	0.064
		median	0.101	0.225	0.051	0.066

(*continued on next page*)

TABLE 8 (continued)

			PreStandard		PostStandard	
			$Bankrupt_{t+1}$		$Bankrupt_{t+1}$	
			No	Yes	No	Yes
	$CostofDebt_{t+1}$	mean	0.077	NA	0.042	NA
		median	0.065	NA	0.037	NA
	$Employees_{t-1,t}$	total	49 to 42	347 to 331	135 to 183	147 to 124
		% neg.	67% (2/3)	75% (3/4)	33% (2/6)	100% (3/3)
	$Employees_{t,t+1}$	total	42 to 56	NA	183 to 173	NA
		% neg.	33% (1/3)	NA	50% (3/6)	NA
GCOpinion	Observations		24 (Type I)	56	30 (Type I)	64
	$Liabilities_t$	mean	2.134	5.570	7.095	27.947
		median	1.120	3.048	2.902	3.446
		total	51.216	311.920	212.850	1,788.608
	$Auditor\Delta_{t,t+\tau}$	percent	13%	NA	30%	NA
	Observations		12 (of 24)	47 (of 56)	20 (of 30)	53 (of 64)
	$\Delta Debt/Assets_t$	mean	0.174	NA	0.010	NA
		median	−0.013	NA	−0.011	NA
	$CostofDebt_t$	mean	0.173	0.115	0.072	0.096
		median	0.119	0.114	0.067	0.071
	$CostofDebt_{t+1}$	mean	0.128	NA	0.083	NA
		median	0.074	NA	0.080	NA
	$Employees_{t-1,t}$	total	233 to 199	3,106 to 2,643	2,332 to 2,349	11,583 to 10,856
		% neg.	27% (4/15)	60% (28/47)	40% (8/20)	51% (27/53)
	$Employees_{t,t+1}$	total	199 to 200	NA	2,349 to 2,061	NA
		% neg.	27% (4/15)	NA	70% (14/20)	NA

Of the 221 stressed, private Belgian companies filing bankruptcy during 1995–1996 and 2001–2002 (see Table 1 for more information), 32 do not have financial statement data to compute *CostofDebt* and another eight do not have employee data for years $t-1$ and t (employee data for year $t+1$ is rarely available for these bankrupt observations). Of the 221 matched non-bankrupt Belgian companies (see Table 1 for more information), 33 do not have financial statement data to compute *CostofDebt* and another 37 do not have employee data for years $t-1$, t, and $t+1$.

Variable Definitions:

Liabilities = total liabilities in millions of 1994 Euros;
AuditorΔ = companies that, at the end of their three-year mandate period, change their audit firm;
ΔDebt/Assets = change in interest-bearing liabilities divided by (prior year) total assets;
CostofDebt = interest expense divided by interest-bearing liabilities (following Pittman and Fortin [2004], this variable is winsorized at +5 percent and −5 percent); and
Employees = total number of employees.

financing and the cost of debt for 12 of the 24 companies that get a GC but do not file bankruptcy the next year. Because of the potential for survivorship and measurement errors, we focus on sample medians.[31] In the pre-standard period, while there is little change in their level of debt financing, the cost of debt for the median company decreases from 11.9 percent to 7.4 percent. In contrast, post-standard, for which we can obtain the necessary data for 20 of 30 Type I errors, the median company's cost of debt increases from 6.7 percent to 8.0 percent. Overall, from the standpoint of the recipient of a Type I error, the effect of the Belgian standard seems somewhat detrimental.

As for the effect of auditor Type I errors on employees, Table 8 provides information on total headcount in the same companies for which we examine debt levels and the cost of debt. In the pre-standard period, during which smaller stressed companies were more likely to receive a GC, there is little evidence of a decrease in overall headcount for companies that get a GC but do not file bankruptcy the next year (i.e., from 199 to 200). In contrast, in the post-standard period, total headcount declines 12 percent (i.e., from 2,349 to 2,061) for companies that receive a GC but do not file bankruptcy the next year, 70 percent of which exhibit a decrease in the number of employees. While we concede that other factors likely contribute to employment changes and other interpretations apply (e.g., workers' councils agree to such reductions to facilitate company survival), from an employee standpoint, the effect of the Belgian standard seems detrimental.

On balance, our results suggest the Belgian GC standard favors creditors and disfavors auditors, and perhaps companies and employees. Whether the benefits outweigh the costs, however, depends on the priorities one assigns to the various parties.

V. SUMMARY AND IMPLICATIONS

In a highly influential study (e.g., the first paper referenced in their 2004 Nobel Prize), Kydland and Prescott (1977, 487) advance the argument for rules rather than discretion by demonstrating the critical importance of making a credible commitment to simple rules:

> In a democratic society, it is probably preferable that selected rules be simple and easily understood, so it is obvious when a policymaker deviates from the policy. There could be institutional arrangements which make it a difficult and time-consuming process to change the policy rules in all but emergency situations.

Boot and Thakor (1993) further advance the case for rules over discretion by arguing that rules help restrain the extent to which self-interest influences a policymaker's decisions.

Drawing on these insights as motivation, we study the relation between the nature of GC audit standards—in terms of whether they promulgate discretion (which, internationally, such standards overwhelming do) or impose rules-based guidance—and auditor reporting on stressed companies. Our context is Belgium where, in 2000, standard-setters increased GC reporting responsibilities by issuing a formal audit standard. In contrast to its predecessor (a non-binding "circular letter" allowing for considerable discretion), the GC standard references simple, credible financial-juridical criteria in Belgian law for board of director GC disclosures. As such, the Belgian GC standard is rules-based, since these criteria influence the extent of the board's GC disclosures and, in turn, the auditor's decision to render a GC opinion.

---

[31] Our concern with measurement error relates to calculating the cost of debt using financial statement information. Following Pittman and Fortin (2004), we compute annual interest expense divided by the average of short- and long-term debt. However, because this approach yields extremely large values for certain companies, we winsorize this measure of annual interest expense at the 5th and 95th percentiles of the opinion-year distribution.

Consistent with the view that the GC standard led to an increase in the costs of auditor acquiescence, for samples of stressed, private bankrupt companies and non-bankrupt companies we report a significant increase in the frequency of GC opinions after the effective date of the GC standard. For bankrupt companies, the increase in GCs provides evidence that, subsequent to the new standard, there is a decrease in the auditor's Type II error. For non-bankrupt companies, the increase in GCs provides evidence of a concomitant increase in the auditor's Type I error. In addition, we also find that the negative relation between company size and the presence of a GC opinion, found in virtually all extant studies, is no longer evident subsequent to the effective date of the Belgian GC standard. That the GC standard results in similar rates of GC opinions for large and small stressed companies suggests greater consistency in auditor reporting. This finding is germane given the International Auditing and Assurance Standards Board's (IAASB) revision of International Standard on Auditing (ISA) 700, *Forming an Opinion and Reporting on Financial Statements*. One of the IAASB's objectives in revising ISA 700 is to promote consistency in the auditor's report (IAASB 2008). Our analysis of the consequences of the decrease (increase) in Type II (Type I) errors suggests that the Belgian standard tends to favor creditors and disfavor auditors, and perhaps also companies and employees. Evaluation of the net of these effects depends on the priorities one assigns to the parties, an evaluation that we respectfully leave to Belgian standard-setters and regulators.

REFERENCES

Aerts, K. 2002. *Tasks and Liabilities of Engagement Partners and Auditors*. Brussels, Belgium: Larcier.

Ashbaugh, H., and T. Warfield. 2003. Audits as a corporate governance mechanism: Evidence from the German audit market. *Journal of International Accounting Research* 2 (1): 1–21.

———, and J. Gassen. 2007. Can audit reforms affect the monitoring role of audits? Evidence from the German market. Working paper, University of Wisconsin–Madison and Humboldt University of Berlin.

Ball, R., and L. Shivakumar. 2005. Earnings quality in UK private firms: Comparative loss recognition timeliness. *Journal of Accounting and Economics* 39 (February): 83–128.

———, and ———. 2008. Earnings quality at initial public offerings. *Journal of Accounting and Economics* 45 (August): 324–349.

Ben-Akiva, M., and S. Lerman. 1985. *Discrete Choice Analysis*. Cambridge, MA: The MIT Press.

Boot, A., and A. Thakor. 1993. Self-interested bank regulation. *The American Economic Review* 83 (May): 206–212.

Carcello, J. V., and Z-V. Palmrose. 1994. Auditor litigation and modified reporting on bankrupt clients. *Journal of Accounting Research* 32 (Supplement): 1–30.

———, D. R. Hermanson, and H. F. Huss. 1995. Temporal changes in bankruptcy-related reporting. *Auditing: A Journal of Practice & Theory* 14 (Fall): 133–143.

———, ———, and ———. 1997. The effect of SAS No. 59: How treatment of the transition period influences results. *Auditing: A Journal of Practice & Theory* 16 (Spring): 114–123.

Carmichael, D., and K. Pany. 1993. Reporting on uncertainties, including going concern. In *The Expectation Gap Standards: Progress, Implementation Issues, Research Opportunities*. New York, NY: AICPA.

Chen, K., and B. Church. 1992. Default on debt obligations and issuance of going-concern opinions. *Auditing: A Journal of Practice & Theory* 11 (Fall): 30–49.

Chaney, P., D. Jeter, and L. Shivakumar. 2004. Self-selection of auditors and audit pricing in private firms. *The Accounting Review* 79 (January): 51–72.

Coffee, J., Jr. 2004. What caused Enron? A capsule social and economic history of the 1990s. *Cornell Law Review* 89 (February): 269–309.

DeAngelo, L. 1981. Auditor size and audit quality. *Journal of Accounting and Economics* 3 (December): 183–199.

DeFond, M., T. Wong, and S. Li. 1999. The impact of improved auditor independence on audit market concentration in China. *Journal of Accounting and Economics* 28 (December): 269–305.

———, K. Raghunandan, and K. R. Subramanyam. 2002. Do non-audit service fees impair auditor independence? Evidence from going-concern audit opinions. *Journal of Accounting Research* 40 (September): 1247–1273.

Dehejia, R., and S. Wahba. 2002. Propensity score-matching methods for nonexperimental causal studies. *The Review of Economics and Statistics* 84 (February): 151–161.

Dewaelheyns, N., and C. Van Hulle. 2008. Legal reform and aggregate small and micro business bankruptcy rates: Evidence from the 1997 Belgian bankruptcy code. *Small Business Economics*.

Dopuch, N., R. Holthausen, and R. Leftwich. 1987. Predicting audit qualifications with financial and market variables. *The Accounting Review* 62 (July): 431–454.

Foster, B., T. Ward, and J. Woodroof. 1998. An analysis of the usefulness of debt defaults and going concern opinions in bankruptcy risk assessment. *Journal of Accounting, Auditing & Finance* 13 (Summer): 351–371.

Francis, J., and J. Krishnan. 2002. Evidence on accounting firm risk-management strategies before and after the Private Securities Litigation Reform Act of 1995. *Asia-Pacific Journal of Accounting and Economics* 9 (December): 135–157.

Friedman, M. 1948. A monetary and fiscal framework for economic stability. *The American Economic Review* 34 (June): 245–264.

Gaeremynck, A., and M. Willekens. 2003. The endogenous relationship between audit-report type and business termination: Evidence on private firms in a non-litigious environment. *Accounting and Business Research* 33 (1): 65–79.

Geiger, M., and K. Raghunandan. 2001. Bankruptcies, audit reports, and the Reform Act. *Auditing: A Journal of Practice & Theory* 20 (March): 187–195.

Gibbins, M., S. Salterio, and A. Webb. 2001. Evidence about auditor-client management negotiation concerning client's financial reporting. *Journal of Accounting Research* 39 (December): 534–563.

Goldman, A., and B. Barlev. 1974. The auditor-firm conflict of interests: Its implications for independence. *The Accounting Review* 49 (October): 707–718.

Hausman, J. 1978. Specification tests in econometrics. *Econometrica* 46 (November): 1251–1271.

Holder-Webb, L., and M. Wilkins. 2000. The incremental information content of SAS No. 59 going-concern opinions. *Journal of Accounting Research* 38 (Spring): 209–219.

Hopwood, W., J. McKeown, and J. Mutchler. 1994. A reexamination of auditor versus model accuracy within the context of the going-concern opinion decision. *Contemporary Accounting Research* 10 (Spring): 409–431.

———, ———, and ———. 1988. The sensitivity of financial distress prediction models to departures from normality. *Contemporary Accounting Research* 5 (Spring): 284–298.

International Auditing and Assurance Standards Board (IAASB). 2008. *Forming an Opinion and Reporting on Financial Statements*. ISA 700. Available at: http://www.ifac.org/IAASB/).

Kida, T. 1980. An investigation into auditors' continuity and related qualification judgments. *Journal of Accounting Research* 18 (Autumn): 506–523.

Knechel, W., and A. Vanstraelen. 2007. The relationship between auditor tenure and audit quality implied by going concern opinions. *Auditing: A Journal of Practice & Theory* 26 (March): 113–131.

Krishnan, J., and J. Krishnan. 1996. The role of economic trade-offs in the audit opinion decision: An empirical analysis. *Journal of Accounting, Auditing & Finance* 11 (Fall): 565–586.

Kydland, F., and E. Prescott. 1977. Rules rather than discretion: The inconsistency of optimal plans. *The Journal of Political Economy* 85 (June): 473–491.

La Porta, R., F. Lopez-de-Silanes, A. Shleifer, and R. Vishny. 1998. Law and finance. *The Journal of Political Economy* 106 (December): 1113–1155.

Louwers, T. 1998. The relation between going-concern opinions and the auditor's loss function. *Journal of Accounting Research* 36 (Spring): 143–156.

Magee, R., and M-C. Tseng. 1990. Audit pricing and independence. *The Accounting Review* 65 (April): 315–336.

Maijoor, S., W. Buijink, and R. Meuwissen. 1998. Towards the establishment of an internal market for audit services within the European Union. *European Accounting Review* 7 (December): 655–673.

Matsumura, E. M., K. R. Subramanyam, and R. Tucker. 1997. Strategic auditor behavior and going-concern decisions. *Journal of Business Finance & Accounting* 24 (July): 727–758.

McKeown, J., J. Mutchler, and W. Hopwood. 1991. Towards an explanation of auditor failure to modify the audit opinions of bankrupt companies. *Auditing: A Journal of Practice & Theory* 10 (Supplement): 1–13.

Mueller, G., H. Gernon, and G. Meek. 1994. *Accounting: An International Perspective*. Burr Ridge, IL: Irwin.

Mutchler, J. 1985. A multivariate analysis of the auditor's going-concern opinion decision. *Journal of Accounting Research* 23 (Autumn): 668–682.

———. 1986. Empirical evidence regarding the auditor's going-concern opinion decision. *Auditing: A Journal of Practice & Theory* 5 (Fall): 148–163.

———, W. Hopwood, and J. McKeown. 1997. The influence of contrary information and mitigating factors on audit opinion decisions on bankrupt companies. *Journal of Accounting Research* 35 (Autumn): 295–310.

Nelson, M., J. Elliott, and R. Tarpley. 2002. Evidence from auditors about managers' and auditors' earnings management decisions. *The Accounting Review* 77 (Supplement): 175–202.

Ng, T., and H.-T. Tan. 2003. Effects of authoritative guidance availability and audit committee effectiveness on auditors' judgments in an auditor-client negotiation context. *The Accounting Review* 78 (July): 801–818.

Ohlson, J. 1980. Financial ratios and the probabilistic prediction of bankruptcy. *Journal of Accounting Research* 18 (Spring): 109–131.

Palmrose, Z-V. 1987. Litigation and independent auditors: The role of business failures and management fraud. *Auditing: A Journal of Practice & Theory* 6 (Spring): 90–103.

Pittman, J., and S. Fortin. 2004. Auditor choice and the cost of debt capital for newly public firms. *Journal of Accounting and Economics* 37 (February): 113–136.

Raghunandan, K., and D. Rama. 1995. Audit reports for companies in financial stress: Before and after SAS No. 59. *Auditing: A Journal of Practice & Theory* 14 (Spring): 50–63.

Reynolds, J. K., and J. R. Francis. 2000. Does size matter? The influence of large clients on office-level auditor reporting decisions. *Journal of Accounting and Economics* 30 (December): 351–373.

Rosenbaum, P., and D. Rubin. 1983. The central role of the propensity score in observational studies for causal effects. *Biometrika* 70 (April): 41–55.

St. Pierre, K., and J. Anderson. 1984. An analysis of the factors associated with lawsuits against public accountants. *The Accounting Review* 59 (April): 242–263.

Vanstraelen, A. 1999. The auditor's going concern opinion decision: A pilot study. *International Journal of Auditing* 3 (March): 41–57.

———. 2002. Auditor economic incentives and going-concern opinions in a limited litigious Continental European business environment: Empirical evidence from Belgium. *Accounting and Business Research* 32 (3): 171–186.

———. 2003. Going-concern opinions, auditor switching, and the self-fulfilling prophecy effect examined in the regulatory context of Belgium. *Journal of Accounting, Auditing & Finance* 18 (Spring): 231–253.

———, and M. Willekens. 2008. Audit regulation in Belgium: Overregulation in a limited capital market oriented country. In *Auditing, Trust and Governance: Developing Regulation in Europe*, edited by R. Quick. London, U.K.: Routledge.

Willekens, M., and C. Achmadi. 2003. Pricing and supplier concentration in the private client segment of the audit market: Market power or competition? *The International Journal of Accounting* 38 (4): 431–455.

———, and A. Gaeremynck. 2005. *Prijszetting in de Belgische Auditmarkt.* (*Price Setting in the Belgian Audit Market.*) Instituut der Bedrijfsrevisoren (Belgian Institute of Registered Auditors) & die Keure.

Wilmots, H. 2004. Het gebruik van de toelichtende paragraaf in het controleverslag (The use of the explanatory paragraph in the auditor's report). *Periodic Newsletter of the Belgian Institute of Registered Auditors* (7).

Winkelman, R. 1997. *Econometric Analysis of Count Data.* Berlin, Germany: Springer.

Wymeersch, E. 1994. Het revisoraat werkt gewoon niet. (Auditing simply doesn't work.) *De Financieel Economische Tijd (The Financial Economic Times)* (October 29).

Zmijewski, M. 1984. Methodological issues related to the estimation of financial stress prediction models. *Journal of Accounting Research* 22 (Supplement): 59–82.

THE ACCOUNTING REVIEW
Vol. 84, No. 5
2009
pp. 1429–1463

American Accounting Association
DOI: 10.2308/accr.2009.84.5.1429

Cross-Listing Audit Fee Premiums: Theory and Evidence

Jong-Hag Choi
Seoul National University

Jeong-Bon Kim
Concordia University and City University of Hong Kong

Xiaohong Liu
The University of Hong Kong

Dan A. Simunic
The University of British Columbia

ABSTRACT: We study the effects of cross-listings on audit fees. We first develop a model in which legal environments play a crucial role in determining the auditor's legal liability. Our model and analysis predict that auditors charge higher fees for firms that are cross-listed in countries with stronger legal regimes than they do for non-cross-listed firms and that the cross-listing audit fee premium increases with the difference in the strength of legal regimes between the cross-listed foreign country and the home country. We then empirically test these predictions. The results of our cross-country regressions strongly support our predictions. In addition, we find no significant cross-listing fee premium for firms that are cross-listed in countries whose legal regimes are

---

We are grateful for the helpful comments from Gary Biddle, Kevin Chen, Tony Kang, Bill Kinney, Sooyoung Kwon, Young Kwon, Jay Junghun Lee, Clive Lennox, Annie Qiu, Jason Schloetzer, Haina Shi, Byron Song, Xiaodong Xu, Yoonseok Zang, and, in particular, from the Steve Kachelmeier (editor) and two anonymous referees. We also thank the participants at the 2005 Korean Accounting Association Conference, the 2006 Annual Conference of Canadian Academic Accounting Association, the 2006 American Accounting Association Annual Meeting, and research workshops at the Hong Kong Polytechnic University, Hong Kong University of Science and Technology, Seoul National University, Singapore Management University, and Xiamen University for their useful comments. We are grateful to Sergei Sarkissian and Michael Schill for sharing their cross-listing data. We appreciate the financial support from Samil PricewaterhouseCoopers in Korea, a postdoctoral fellowship grant at the Hong Kong Polytechnic University, the seed funding for basic research at the University of Hong Kong, the Social Sciences & Humanities Research Council of Canada, and the KPMG Research Bureau in Financial Reporting at UBC. All errors are our own.

Editor's note: Accepted by Steven Kachelmeier, with thanks to Bill Kinney for serving as editor on previous versions.

Submitted: May 2006
Accepted: January 2009
Published Online: September 2009

no stronger than those of their home countries. This suggests that cross-listing audit fee premiums are associated with increased legal liability and not with increased audit complexity *per se*. Our findings help explain why cross-listing premiums occur and what determines their magnitude.

Keywords: *cross-listing; audit fee; legal regime; audit quality.*

Data Availability: *Data are available from public sources identified in the study.*

I. INTRODUCTION

Cross-listing is an important, strategic decision for firms that are raising capital. It typically causes a shift in the legal liability and enforcement regime, which, in turn, affects the interests of existing shareholders, potential new investors, auditors, and other stakeholders in the firm. This study investigates how the shareholders' decision to cross-list the firm's shares in a foreign capital market affects audit pricing. In particular, we are interested in the following questions: Do auditors in a home country charge higher audit fees for cross-listed firms (CL firms) than they do for non-cross-listed firms (NCL firms)? Do they charge higher audit fees for all CL firms or only for those CL firms that are listed in foreign countries with stronger legal regimes? Stated differently, is the change in the auditor's legal liability the driving force behind any fee premiums (which we call cross-listing audit fee premiums) charged for CL firms relative to NCL firms? By addressing these questions, we seek to offer insights into the effects of cross-listings and the associated change in legal environments on the auditor's behavior and audit fees.

We first develop a simple model in which the strength (or strictness) of a country's legal regime plays a crucial role in determining the auditor's legal liability. For CL firms, in the event of audit failure, the stronger of the legal regimes, i.e., that of the home country or that of the foreign country, determines whether the auditor is negligent and thus liable for client damages that arise from that audit failure (Stulz 1999). In a competitive audit market, the audit fee equals the sum of the expected legal liability cost and the auditor's effort cost that is associated with audit quality and audit complexity.

Based on our analysis of the effects of cross-listings on audit fees, we formulate the following three hypotheses. First, when firms cross-list in a country with a stronger legal regime than that of their home country, audit fees are higher for the CL firms than they are for NCL firms. The reason is that an increase in the strength of a legal regime raises the expected legal liability cost, which, in turn, raises the audit fee. Second, when firms cross-list in a country with the same or a weaker legal regime than that of their home country, audit fees are higher for the CL firms than they are for NCL firms if, and only if, cross-listings require more complex audits. Third, for firms cross-listed in foreign countries with stronger legal regimes, the cross-listing fee premiums increase with the difference in the strength of legal regimes between the cross-listed foreign countries and the CL firm's home countries.

In addition, we show that a stronger legal regime induces higher audit quality, and that the value of an audit (reflected by improved investment decisions) that is associated with cross-listing in a stronger legal regime increases with the amount of a firm's new investment. Thus, a firm with a greater demand for investment funding is more likely to benefit from cross-listing in a stronger legal regime, despite the audit fee premium and other costs associated with the cross-listing. We thus offer insights into the value of auditing in a cross-listing context.

We test our three hypotheses using audit fee data from 14 countries with different legal regimes. The results of various tests strongly support the first and third hypotheses. Furthermore, we find no significant audit fee premium for firms that are cross-listed in countries whose legal regimes are weaker than or similar to those of their home countries. According to the second hypothesis, this evidence suggests that cross-listings do not significantly increase audit complexity, and thus the observed cross-listing fee premiums are primarily driven by the increase in legal liability costs associated with stronger legal regimes.

Cross-listing is often viewed as a credible commitment on the part of shareholders to submit their firms to a stronger legal regime.[1] Prior studies have documented various cross-listing benefits. Specifically, they find that cross-listing on the U.S. stock exchanges leads to greater analyst following or better information environments (Lang, Lins, and Miller 2003), higher quality accounting information (Lang, Raedy, and Yetman 2003), wider investor recognition (Foerster and Karolyi 1999), lower costs of capital (Miller 1999; Lee 2003; Hail and Leuz 2008), stronger investor protection and thus higher market valuation (Stulz 1999; Reese and Weisbach 2002; Doidge et al. 2004),[2] and more voluntary disclosures (Shi et al. 2008).

However, this literature (except for Seetharaman et al. [2002] discussed below) has paid little attention to the effect of cross-listings on the value of audits and audit fees. Our study contributes to the cross-listing literature by adding a carefully developed auditing perspective to that literature. Moreover, prior studies focus on cross-listings in the U.S. market (which has the strongest legal regime), while our study considers both cross-listings in strong legal regimes and those in weak legal regimes. We examine how the cross-listing effects differ across different legal regimes and thus provide broader evidence than previous U.S.-based studies.

Our study is related to that of Choi et al. (2008) in the sense that both investigate the role of legal regimes in audit pricing. That study focuses on the impact of legal regimes on Big 4 auditor fee premiums, whereas this study focuses on the impact of cross-listings and the associated change in legal regimes on audit fees paid by CL versus NCL firms. Our study is also related to that of Seetharaman et al. (2002), who find that auditors in the U.K. charge audit fee premiums for clients that are cross-listed in the U.S., but not for those cross-listed in non-U.S. countries.[3] However, unlike our study using cross-listing data from 14 countries, their single-country (U.K.) setting is unable to show or explain how the cross-listing fee premium changes with the legal regimes of the home and cross-listed foreign countries.

Section II presents the model, the analytical results, and the hypotheses. Section III explains the empirical procedures that are used to test the hypotheses. Section IV describes the sample and data and presents the descriptive statistics. Section V presents the regression results. Section VI presents the results of various sensitivity tests, and Section VII concludes.

---

[1] Other mechanisms that may be used to protect the interests of investors (in countries with weak legal regimes) include good corporate governance and high-quality auditors, i.e., Big 4 auditors (Choi and Wong 2007; Fan and Wong 2005; Klapper and Love 2004).

[2] Although this line of research focuses on one type of agency problem (i.e., the expropriation of small shareholders by large shareholders), we focus on the other type of agency problem, i.e., the poor investment decisions made by shareholders (due to the information asymmetry between managers and shareholders) that can be mitigated by a high-quality audit.

[3] An indicator variable is used to represent non-U.S. countries in the regressions of their study.

II. MODEL AND HYPOTHESES

Consider a one-period setting in which the continued operation of a firm requires an investment of $I > 0$ at the beginning of the period. The firm is owned by its shareholders and is operated by a manager. The productivity of the firm depends on the manager's competence, which is either high or low. The firm's shareholders do not know the manager's type. *A priori*, the manager is of the high type with probability $p \in (0, 1)$ and of the low type with probability $(1 - p)$. The firm will generate a cash flow of R_H with a high-type manager, but R_L with a low-type manager at the end of the period. We assume that $pR_H + (1 - p)R_L > I$ and $R_L < I$. Thus, the firm is worth investing in without knowledge of the manager's type, but not if the manager is known to be of the low type. We assume that although a correctly stated financial report can reveal the manager's competence, the manager is able to obtain personal benefits from the continued operation of the firm and thus always presents financial statements that implicitly report a high level of competence.[4] This means that the financial report issued by a high-type manager is correctly stated, but the report by a low-type manager is misstated. The existing shareholders are cash flow constrained and thus need to raise the required funds by selling part of their ownership in the firm. For simplicity, we assume that if the shareholders do not attract new investments, they will liquidate the firm and get a zero return.[5] All of the parties in the model are risk-neutral and the interest rate is normalized to zero.

The shareholders can hire an auditor to attest to the correctness of the manager's report. The more effort an auditor exerts, the more likely he is to detect the misstatements in a (misstated) financial report. Formally, let $q \in [0,1]$ be audit quality and kq be the auditor's effort cost, where $k > 0$. After conducting the audit, the auditor has evidence as to the correctness of the financial report. In this one-period model, there is no incentive for the auditor to misreport this evidence. Thus, when the auditor obtains no evidence of misstatements, he will issue an unqualified opinion; when he obtains evidence of misstatements, he will issue a qualified opinion. As is common in the literature (e.g., Dye 1995; Pae and Yoo 2001), audit technology is characterized by the probability of an auditor issuing a correct opinion about a financial report as follows:

$$\text{Prob}(\text{unqualified}|\text{High}, q) = 1 \text{ and } \text{Prob}(\text{qualified}|\text{Low}, q) = q. \tag{1}$$

High (Low) indicates that the financial report is prepared by a high- (low-) type manager. Recall that the financial report of a high-type manager is correctly stated, and that of a low-type manager contains material misstatements. Thus, Expression (1) means that an auditor will not find any misstatements if the manager's report contains none, but may fail to detect misstatements if this report is misstated.

Cash flow is realized, and thus the correctness of the manager's report is revealed at the end of the period. Audit failure occurs when the auditor issues an unqualified opinion on a misstated report. We assume that in the event of audit failure, the probability of the auditor making a legal payment (i.e., client damages and penalties decided by the court) to the client (i.e., all of the shareholders) is $r(1 - q)$, where $r \in (0, 1]$ represents the strength or strictness of the legal liability regime. We can interpret $r(1 - q)$ as the probability that in the event of audit failure, a court decides that the auditor is negligent and thus responsible for the client damages arising from that failure. Note that the stronger is the legal regime, r (or the lower is audit quality, q), the more likely it is that audit failure will eventually lead to the auditor making a legal payment to the client. Let this legal payment be denoted by $L > 0$.[6]

---

[4] Such personal benefits may include the business experience/skills a manager gains from operating a firm, the cash flows he diverts from the firm, and other job-related advantages. The agency problem between the shareholders and the manager is assumed to be exogenous in this simple model.

[5] Alternatively, we could assume that the liquidation value of the firm is nonzero and varies with the manager's type, but this would complicate the notation without qualitatively affecting our analyses.

The existing shareholders sell a fraction of the ownership of the firm, denoted by λ, to raise the necessary investment, I. The capital market is competitive so that new investors are willing to invest if their expected payoff is zero. For comparison purposes, first consider a setting without auditing. The market value of a firm, denoted by V_0, equals the expected cash flow of that firm: $V_0 = pR_H + (1 - p)R_L$. As the expected payoff to the new investors is zero, we have $\lambda_0 V_0 - I = 0$. The existing shareholders' payoff is $\pi_0 = (1 - \lambda_0)V_0 = V_0 - I$.

Now, consider a setting in which the existing shareholders pay an audit fee, denoted by f, to hire an auditor to verify the manager's report. The investors will not invest when a qualified opinion is issued, as $R_L < I$. Let $t \equiv$ Prob)High|unqualified, q) $= \frac{p}{p+(1-p)(1-q)}$. Given an unqualified opinion, the market value of the firm becomes $V = tR_H + (1 - t)[R_L + r(1 - q)L]$. As we will show later, the higher the audit quality, the greater the market value of the firm. The fraction of ownership sold to the new investors is $\lambda = \frac{I}{V}$, which decreases with the firm's value. The existing shareholders' payoff is $\pi = \text{Prob(unqualified}|q)(1 - \lambda)V - f$.

These shareholders may choose to cross-list (sell) the firm's shares in a foreign country. Let r_h and r_f denote the strength of the legal regimes in the home country and the foreign country, respectively. For a CL firm, in the event of audit failure, the stronger of the two legal regimes, i.e., that of the home country or that of the foreign country, determines whether the auditor is negligent and liable for the audit failure.[7] Let r_c denote the strength of the applicable legal regime for a CL firm; $r_c = \max\{r_h, r_f\}$. The audit of a CL firm may involve greater complexity than that of a NCL firm; we therefore assume that to achieve a given level of audit quality, the auditor's effort costs (e.g., audit hours) for a CL firm are no less than they are for a NCL firm: $k_c q \geq k_h q$, namely, $k_c \geq k_h$. In this regard, k can be interpreted as a parameter that captures the complexity of the audit.

The auditor's decision problem is to choose a level of audit quality, q, to minimize the expected audit cost, that is:

$$\underset{q\in[0,1]}{\text{Minimize}}\ (1 - p)(1 - q)[r(1 - q)]L + kq. \tag{2}$$

The first term in objective function (2) is the auditor's expected legal liability cost and the second term is the auditor's effort cost. The audit quality, q, reduces the expected liability cost by reducing both the probability of an audit failure, $(1 - q)$, and the probability of the auditor's being found negligent and liable in court in the event of an audit failure, $[r(1 - q)]$.

---

[6] To keep our analysis focused, the model does not distinguish between Big 4 and non-Big 4 auditors, but our analytical results about the incentive effects of cross-listings apply to both the Big 4 and non-Big 4 auditors.

[7] For example, Stulz (1999) argues that firms may "rent" foreign governance and legal systems by cross-listing in countries in which minority protection rights are stronger. Although there have been no empirical studies on cross-country lawsuits involving auditors, anecdotal evidence exists. For example, the Italian audit firm, Deloitte and Dianthus SpA, was ordered by a U.S. court to make legal settlement to Parmalat (an Italian dairy company) to cover investor claims that accused the audit firm of helping former Parmalat management to overstate accounting reports. Parmalat was traded in the U.S. OTC market until its collapse in 2003 (see http://www.accountancyage.com/accountancyage/news/2183799/deloitte-ordered-pay-66m).

To ensure an interior solution, we assume that $L > \frac{k}{2r(1-p)}$. The optimal level of audit quality denoted by q^* (derived from the first order condition) is:

$$q^* = \frac{2(1-p)rL - k}{2(1-p)rL} \in (0,1). \tag{3}$$

Consistent with intuition, the optimal audit quality, q^*, increases with the strength of a legal regime ($\partial q^*/\partial r > 0$), but decreases with audit complexity ($\partial q^*/\partial k < 0$).

Assuming that the market for audit services is competitive so that the auditor gets a zero payoff in equilibrium, we can express the equilibrium audit fee as:

$$f = (1-p)r(1-q^*)^2 L + kq^* \tag{4}$$

which represents the expected legal liability cost plus the effort costs in equilibrium. Substituting Equation (3) into Equation (4), we can write the audit fee as a function of the strength of legal regime r and the audit complexity k as:

$$f(r,k) = k - \frac{k^2}{4(1-p)rL}. \tag{5}$$

Let $\Delta_r = r_c - r_h$ and $\Delta k = k_c - k_h \geq 0$. To examine how a cross-listing affects the audit fee, we derive:

$$\Delta f = f(r_c,k_c) - f(r_h,k_h) = \left.\frac{\partial f}{\partial r}\right|_{(r_h+\theta_1\Delta r, k_h)} \Delta r + \left.\frac{\partial f}{\partial k}\right|_{(r_c, k_h+\theta_2\Delta k)} \Delta k \tag{6}$$

where θ_1 and θ_2 are constants between 0 and 1 (see Widder [1989] for details.) As $\frac{\partial f}{\partial r} > 0$, $\frac{\partial f}{\partial k} > 0$ (as shown by Equations (7) and (8) below), and $\Delta k \geq 0$, we know from Equation (6) that if $\Delta r > 0$, then $\Delta f > 0$, and that if $\Delta r = 0$ and $\Delta k > 0$ ($\Delta k = 0$), then $\Delta f > 0$ ($\Delta f = 0$). This yields the following observation.

Observation 1: The audit fee is higher for a firm that cross-lists its shares in a country with a stronger legal regime, compared with a non-cross-listed firm. The audit fee is higher (the same) for a firm that cross-lists its shares in a country with the same or a weaker legal regime if $\Delta k > 0$ (if $\Delta k = 0$), compared with a non-cross-listed firm.

The reason that an audit fee premium is paid by a firm that cross-lists in a stronger legal regime is as follows. The change in the audit fee, Δf, consists of two parts, as illustrated by Equation (6). First, with audit complexity k being fixed, the increase in the strength of a legal regime has a positive impact on the audit fee, that is:

$$\frac{\partial f}{\partial r} = (1 - p)(1 - q^*)^2 L > 0. \tag{7}$$

To explain, note that an increase in the strength of the legal regime raises the expected-legal-liability-cost component of the audit fee, which, in turn, raises the audit fee to a higher level.[8]

Second, with the legal regime fixed, a potential increase in the complexity of the audit also has a positive impact on the audit fee, that is:

$$\frac{\partial f}{\partial k} = q^* > 0. \tag{8}$$

Note here that audit complexity increases the audit fee through its positive effect on the auditor's effort cost. In sum, the audit fee premiums for firms that are cross-listed in stronger legal regimes could be due to both the increase in the strength of the legal regime and the increase in audit complexity.

Observation 1 also states that when a firm is cross-listed in a country with a legal regime that is similar to or weaker than that of the home-country (e.g., a Hong Kong firm cross-listed in the U.K.), the necessary (and sufficient) condition for observing a cross-listing fee premium is that a cross-listing increases audit complexity ($\Delta k > 0$). We formulate the following two hypotheses based on Observation 1.

H1: Given that a firm cross-lists in a country with a stronger legal regime than that of its home country, i.e., $r_f > r_h$, the audit fee is higher for the cross-listed firm than it is for a non-cross-listed firm.

H2: Given that a firm cross-lists in a country with the same or a weaker legal regime than that of its home country, i.e., $r_f \leq r_h$, the audit fee is higher for the cross-listed firm than it is for a non-cross-listed firm if, and only if, the audit of a cross-listed firm involves greater audit complexity, i.e., $k_c > k_h$.

We next examine how the cross-listing audit fee premiums change with the difference in the strength of legal regimes between the cross-listed foreign country and the home country. We obtain that:

$$\frac{\partial \Delta f}{\partial \Delta r} = \frac{k_c^2}{4(1 - p)r_c^2 L} > 0, \tag{9}$$

which leads to the following observation.

Observation 2: The audit fee premium that is associated with a firm that cross-lists in a foreign country with a stronger legal regime increases with the difference in the strength of legal regimes between the foreign country and the home country.

---

[8] The effect of an increase in the strength of the legal regime on the audit fee is not as straightforward as it appears. On the one hand, with audit quality being fixed, an increase in the strength of the legal regime will increase auditor legal liability, which will increase the audit fee. On the other hand, an increase in the strength of the legal regime will induce higher audit quality, which will then reduce the audit failure rate (represented by $1 - q$) and thus have a negative effect on the expected-legal-liability component of the audit fee. A determination of which of the above opposite effects on the audit fee dominates involves the Envelope Theorem.

This result is intuitive. Recall that the audit fee increases with the strength of a legal regime. Auditors will thus charge higher fees for firms that are cross-listed in countries with stronger legal regimes than they will for firms that are cross-listed in countries with weaker legal regimes. As a result, the cross-listing fee premium becomes larger as the home country's legal regime deviates further from that of the foreign country. Hypothesis 3 follows immediately from Observation 2.

> **H3:** For a given strong legal regime in the foreign country in which a firm is cross-listed, the cross-listing audit fee premium is larger (smaller) when the legal regime of the firm's home country is weaker (stronger).

In equilibrium, the value of an audit is $\pi - \pi_0 = [(1 - p)(I - R_L) - k]q^*$, which represents the improved investment decisions as a result of an audit. To ensure that the value of an audit is positive (which also ensures the existence of an audit market), we assume that $(1 - p)(I - R_L) > k$, which means that the wasteful investments saved by an audit of the firm with a low-type manager is greater than the marginal cost of audit quality. Clearly, the value of an audit increases with the quality of the audit, $\frac{\partial(\pi - \pi_0)}{\partial q} > 0$. In addition, note that $\frac{\partial V}{\partial q} > 0$ (which is inferred from $\frac{\partial(\pi - \pi_0)}{\partial q} = -(1 - p)(V - I) + [p + (1 - p)(1 - q)]\frac{\partial V}{\partial q} > 0$). This means that the higher the audit quality, the greater the market value of the firm. In other words, a stronger legal regime induces a higher audit quality, which allows the existing shareholders to sell their shares in the ownership of the firm to new investors at a higher market price. This illustrates the benefits of cross-listings generated by higher quality audits.

To capture the situation in which cross-listing in a stronger legal regime often involves significant compliance costs (such as complying with the more rigorous regulatory requirements of the foreign country), we assume that the existing shareholders pay a cost, denoted by w, if they cross-list their firm in a stronger legal regime. The benefit to the shareholders from the cross-listing in a stronger legal regime is thus:

$$\pi_c - \pi_h = (1 - p)(I - R_L)[q_c^* - q_h^*] - (k_c q_c^* - k_h q_h^*) - w,$$

which equals the incremental value of the audit that is due to the higher audit quality induced by the stronger foreign legal regime, minus the incremental audit effort cost, and minus the compliance cost. It is clear that the shareholders' cross-listing benefit increases with the level of a firm's investment I, as $q_c^* > q_h^*$.

> **Observation 3:** For any given cost to existing shareholders, w, that is associated with cross-listing in a country with a stronger legal regime, there exists an investment threshold, denoted by I^*, such that the shareholders are better off by cross-listing their firm in that country when the required investment for the firm exceeds the threshold, i.e., when $I > I^*$. The threshold is $I^* = R_L + \frac{w + k_c q_c^* - k_h q_h^*}{(1 - p)(q_c^* - q_h^*)}$.

This observation implies that firms with greater funding requirements are more likely to cross-list in stronger legal regimes to allow their shareholders to enjoy the higher value of audits in spite of the compliance costs and increased audit fees. Thus, our analysis can explain both the shareholders' decision to submit their firm to a stronger legal regime through a cross-listing and the value of auditing in this context. This result is also consistent with the evidence from Doidge et al. (2004) that firms that are cross-listed in the U.S. have greater growth opportunities (as reflected by their higher Tobin's Q ratios) compared with non-cross-listed firms.

III. EMPIRICAL PROCEDURES

Research Design for Testing H1 and H2

We use the following regression model to test H1 and H2:

$$\begin{aligned} AUDFEE = {} & \beta_0 + \beta_1 REGIME + \beta_2 CROSS + \sum_{k=1}^{8} \delta_k FSCONTROL \\ & + \sum_{k=1}^{4} \varphi_k CLCONTROL_k + \text{Fixed Effects} + \text{Error Term} \end{aligned} \tag{10}$$

where *FSCONTROL* and *CLCONTROL* denote firm-specific control variables and country-level control variables, respectively. All of the variables are defined in Table 1. We include eight firm-specific controls (*LNTA*, *INVREC*, *LOSS*, *LEV*, *ISSUE*, *NBS*, *NGS*, and *BIG4*) and four country-level controls (*DISCL*, *B4DO*, *GDP*, and *FDI*). As shown in Table 1, *AUDFEE*, *CROSS*, and the eight firm-specific control variables are measured at the firm level for each country in each year; *REGIME* and *DISCL* are measured for each country with no variation over years; *B4DO*, *GDP*, and *FDI* are measured for each country in each year.

Our test variable, *CROSS*, is an indicator variable that equals 1 for CL firms, and 0 otherwise. The coefficient on *CROSS* (β_2) thus captures the cross-listing audit fee premium (i.e., the difference in audit fees between CL and NCL firms). We identify cross-listed firms based on the foreign equity trading data retrieved from Worldscope. The other variable, *REGIME*, represents the strength of a home country's legal regime. The coefficient on *REGIME* (β_1) captures the impact of the home country's litigation environment on the audit fee. We expect a positive β_1 as the audit fee increases with the strength of a legal regime. *REGIME* is defined as the natural log of the Wingate (1997) litigation index, which was developed by an international insurance underwriter for one of the Big 4 audit firms based on an assessment of the litigation risk of doing business as an auditor in each country. The index takes into account a country's legal, regulatory, political, and economic environments. Other studies that use the Wingate index as a proxy for a country's litigation risk or legal environment include Choi and Wong (2007), Choi et al. (2008), Chung et al. (2008), and Francis et al. (2003).

We include eight firm-specific control variables. *LNTA* and *INVREC* are proxies for client size and client complexity, respectively (e.g., Simunic 1980; Francis 1984). We include *NBS* and *NGS* as additional proxies for client complexity as operationally or geographically diversified firms may require more complex audits (e.g., Simunic 1980; Choi et al. 2008). Similar to Simunic (1980) and Seetharaman et al. (2002), we include *LOSS* and *LEV* to measure the client-specific litigation risks to be borne by auditors. The demand for quality audits and the audit risk are likely to be high when firms are involved in equity or bond offerings (Choi and Wong 2007); hence, we include *ISSUE* to capture these potential effects on audit pricing. We also include a Big 4 indicator variable (*BIG4*) to capture a Big 4 auditor fee premium (e.g., DeFond et al. 2000; Francis and Simon 1987). As audit

TABLE 1
Empirical Definitions of Variables

Variable	Empirical Definition	Data Source
The Dependent Variable and Test Variables for Firm *i* in Country *j* in Year *t*		
$AUDFEE_{ijt} =$	natural log of audit fee in thousands of U.S. dollars;	Worldscope
$REGIME_{j} =$	the strength of a country's legal regime, measured by the natural log of Wingate's (1997) litigation index; and	Wingate (1997)
$CROSS_{ijt} =$	1 when a firm is cross-listed in a foreign country, and 0 otherwise.	Worldscope
Firm-Specific Control Variables (*FSCONTROL*) for Firm *i* in Country *j* in Year *t*		
$LNTA_{ijt} =$	natural log of year-end total assets in thousands of U.S. dollars;	Worldscope
$INVREC_{ijt} =$	the sum of inventories and receivables divided by total assets;	Worldscope
$LOSS_{ijt} =$	1 when a firm reports a net loss, and 0 otherwise;	Worldscope
$LEV_{ijt} =$	the ratio of year-end total liabilities to total assets;	Worldscope
$ISSUE_{ijt} =$	1 when a firm has obtained long-term debt or equity financing by issuing bonds or new shares to outside capital suppliers in the three years prior to year *t*, and 0 otherwise;	Worldscope
$NBS_{ijt} =$	natural log of 1 plus the number of business segments;	Worldscope
$NGS_{ijt} =$	natural log of 1 plus the number of geographical segments; and	Worldscope
$BIG4_{ijt} =$	1 when a firm uses one of the Big 4 (5, 6, or 8) auditors, and 0 otherwise.	Worldscope
Country-Level Control Variables (*CLCONTROL*) for Country *j* in Year *t*		
$DISCL_{j} =$	a country's disclosure level measured by the CIFAR index (adjusted for the median value of the index);	CIFAR
$B4DO_{jt} =$	Big 4 dominance in a country, measured by the difference between the market share of the smallest Big 4 auditor and that of the largest non-Big 4 auditor;	Worldscope
$GDP_{jt} =$	gross domestic product per capita in thousands of U.S. dollars; and	Int'l Fin'l Statistics by the IMF
$FDI_{jt} =$	foreign direct investment scaled by total GDP.	Int'l Fin'l Statistics by the IMF

fees are normally positively related to client size, client complexity, client-specific risk factors, and auditor quality, we expect all of the coefficients from δ_1 to δ_8 to be positive.

We include four country-level control variables that may affect cross-country variations in audit fees. First, *DISCL* represents a country's disclosure level, as measured by the disclosure index developed by the Center for International Financial Analysis and Research (CIFAR). We use the median-adjusted value of *DISCL* (i.e., the raw value minus the median value of 79) in our regression analyses. A higher level of disclosure may entail more audit work, thus suggesting a positive coefficient on this variable.[9] Second, *B4DO* represents the

---

[9] Instead of the CIFAR index, which is based on actual disclosure practice, we also use the index of mandatory disclosure requirements, which was developed by La Porta et al. (2006; hereafter, LLS). The use of the LLS index does not alter our statistical inferences on the variables of interest, i.e., *CROSS* and *CROSS * REGIME*. We also find that, contrary to our expectation, India and Malaysia rank high under the LLS index, although both rank low under the CIFAR index.

Big 4 auditors' dominance in a country. We compute *B4DO* by taking the difference between the market share of the smallest Big 4 auditor and that of the largest non-Big 4 auditor in the country. The data required to compute *B4DO* are from Worldscope. As the monopoly power of the Big 4 auditors may allow them to charge higher fees, we expect a positive coefficient on *B4DO*. Third, we expect *GDP* (gross domestic product) to have a positive coefficient, as audit fees are likely to be higher in rich countries than in poor countries. Fourth, the demand for audit services is likely to be greater in countries with more foreign direct investments (*FDI*) than in countries with less; thus, we expect a positive coefficient on *FDI*. We extract the data on *GDP* and *FDI* from *International Financial Statistics*, which is published annually by the International Monetary Fund. Finally, we include fixed effect indicator variables for countries and years to control for potential variations in audit fees across countries and over time.

To test H1 and H2, we partition the full sample into two subsamples: (1) the sample of CL firms that are listed in foreign countries whose legal regimes are stronger than those of the home countries and NCL firms in the home countries, which is referred to as subsample H1; and (2) the sample of CL firms that are listed in foreign countries whose legal regimes are no stronger than (i.e., similar to or weaker than) those of the home countries and NCL firms in the home countries, which is referred to as subsample H2. We then estimate Regression (10) separately for each subsample.

In the regression on subsample H1, we expect the coefficient on *CROSS* to be positive ($\beta_2 > 0$), as H1 states that cross-listing in a stronger legal regime increases the audit fee. Hypothesis 2 states that when a cross-listing does not submit a firm to a stronger legal regime, one would observe a significantly positive coefficient on *CROSS* if, and only if, a cross-listing increases audit complexity (e.g., if a CL firm has more complex financial statements than a NCL firm). Hence, a significantly positive β_2 in the regression on subsample H2 would suggest that cross-listings increase audit complexity, whereas an insignificant β_2 would suggest that cross-listings do not necessarily increase audit complexity.

Research Design for Testing H3

Hypothesis 3 states that the fee premiums for firms cross-listed in a given stronger legal regime (e.g., the U.S. with a Wingate index of 15) are larger for the firms that come from countries with relatively weak legal regimes (e.g., Pakistan with a Wingate index of 1) than they are for those from countries with relatively strong legal regimes (e.g., Denmark with a Wingate index of 4.82). To test H3, we run the following regression in which we add the interaction, *CROSS * REGIME,* to Equation (10):

$$\begin{aligned} AUDFEE = {} & \beta_0 + \beta_1 REGIME + \beta_2 CROSS + \beta_3 CROSS * REGIME \\ & + \sum_{k=1}^{8} \delta_k FSCONTROL + \sum_{k=1}^{4} \varphi_k CLCONTROL_k \qquad (11) \\ & + \text{Fixed Effects} + \text{Error Term}. \end{aligned}$$

We fix the strength of the foreign legal regime in which firms are cross-listed and allow only the legal regimes of the home countries to vary. For this purpose, we choose one of the two benchmarks as the cross-listed foreign regime: the U.S., which has the highest Wingate index of 15 (or $REGIME = 2.71$), and the four countries (Australia, Hong Kong, New Zealand, and the U.K.) that have the second-highest Wingate index of 10 (or $REGIME = 2.3$). We thus estimate Equation (11) separately with two data sets: (1) the firms that are cross-listed in the U.S. and NCL firms in their home countries; and (2) the firms that are cross-listed in any of the above four countries (but not cross-listed in the U.S. at the same time) and NCL firms in their home countries.

Hypothesis 3 would be supported if the coefficient on *CROSS* * *REGIME* is significantly negative ($\beta_3 < 0$), which means that the cross-listing fee premium captured by ($\beta_2 + \beta_3$ * *REGIME*) becomes smaller as the CL firms' home-country legal regime (as measured by *REGIME*) becomes stronger.

IV. SAMPLE AND DESCRIPTIVE STATISTICS

Our initial sample consists of all firms for which audit fee data are available from the 2004 Worldscope database during the seven-year period 1996–2002. We choose 1996 as the start of the sample period because, prior to 1996, audit fee data are available only for a small number of the firms covered by Worldscope. After imposing the data requirements to compute the firm-specific variables in Equation (10), we obtain 17,837 firm-year observations from 14 non-U.S. countries (see Column (1), Table 2). We exclude U.S. firms as very few of them (only four out of 3,722 U.S. firm-year observations) are cross-listed in foreign countries.

For ease of notation, we also refer to firm-year observations as firms. Table 2 presents descriptive statistics on cross-listings. Columns (2) and (3) report the numbers of firms that are cross-listed in just one country and those cross-listed in two or more countries. Columns (4) and (5) report the total number of cross-listings and their percentage relative to the sample firms in each country. While the average percentage of CL firms in a country is 10.74 percent, it varies from about 4 percent in Denmark to about 84 percent in Ireland.[10] Column (6) reports the number of firms that are cross-listed in foreign countries whose legal regimes are no stronger than those of their home countries, whereas Column (7) reports the number of firms that are cross-listed in stronger legal regimes. As expected, the majority of the CL firms cross-list in stronger legal regimes.[11] Column (8) reports the number of firms that are cross-listed in the U.S. We find that out of the total of 1,916 CL firms, more than half (980 firms from 13 countries) are cross-listed in the U.S.[12]

Table 3 presents the mean values of the variables included in Equation (10) for each country and the grand mean for the 14 countries in our sample. We also partition the sample countries into two groups: (1) the strong-home-regime sample of seven countries with the Wingate index greater than or equal to 6.22 (Australia, Hong Kong, Ireland, Italy, New Zealand, Norway, and the U.K.); and (2) the weak-home-regime sample of seven countries with the Wingate index less than 6.22 (Denmark, India, Malaysia, Pakistan, Singapore, South Africa, and Sweden).[13] We then perform t-tests for the mean differences between the two groups.

With respect to the test results presented in the last row of Table 3, the following are noteworthy. First, our measure of audit fee (*AUDFEE*) is, on average, significantly higher for the strong-home-regime sample than it is for the weak-home-regime sample. Second, the firms in the strong-home-regime sample are, on average, larger in size (*LNTA*), have

---

[10] Of the 266 cross-listed Irish firms, 238 are cross-listed in the U.K. The results of the tests with and without the Irish firms are qualitatively similar.

[11] There are a few observations that are cross-listed in both weaker (or similar) regime countries and stronger regime countries. We classify them as being cross-listed in the stronger regime countries because the cross-listing in the stronger legal regime may dominate the audit fee effect. The empirical results are almost identical with or without these few observations.

[12] Table 2 reveals that the majority of the 745 firms that are cross-listed in stronger legal regimes other than the U.S. come from two countries: 203 from Ireland and 436 from Malaysia. Further analyses reveal that these 639 firms are mostly cross-listed in the U.K.

[13] The results using the Wingate index of 10 as an alternative cut-off score are qualitatively similar.

TABLE 2
Cross-Listing Statistics

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Country	Total Number of Firm-Year Observations	Number of Observations Cross-Listed in One Country	Number of Observations Cross-Listed in Two or More Countries	Total Number of Cross-Listed Observations [= (2) + (3)]	Percentage of Cross-Listed Observations [= (4)/(1)]	Number of Observations Cross-Listed in Weaker or Similar Regimes	Number of Observations Cross-Listed in Stronger Regimes	Number of Observations Cross-Listed in the U.S.
Australia	913	150	4	154	16.87	69	85	85
Denmark	697	20	7	27	3.87	0	27	20
Hong Kong	2,657	176	2	178	6.70	23	155	155
India	885	96	0	96	10.85	0	96	96
Ireland	316	266	0	266	84.18	0	266	63
Italy	59	31	1	32	54.24	0	32	32
Malaysia	2,199	461	0	461	20.96	2	459	23
New Zealand	330	20	0	20	6.06	5	15	15
Norway	418	21	12	33	7.89	7	26	14
Pakistan	55	5	0	5	9.09	0	5	0
Singapore	1,556	117	7	124	7.97	7	117	84
South Africa	626	81	9	90	14.38	0	90	50
Sweden	370	23	12	35	9.46	0	35	26
U.K.	6,756	341	54	395	5.85	78	317	317
Mean (Total)	(17,837)	(1,808)	(108)	(1,916)	10.74	(191)	(1,725)	(980)

TABLE 3
Sample Characteristics

Country	n	*AUDFEE* (mean)	*Wingate Index*	*LNTA* (mean)	*INVREC* (mean)	*LOSS* (mean)	*LEV* (mean)	*ISSUE* (mean)	*NBS* (mean)	*NGS* (mean)	*BIG4* (mean)	*Raw DISCL*	*B4DO* (mean)	*GDP* (mean)	*FDI* (mean)
Australia	913	5.18	10	11.64	0.19	0.28	0.48	0.84	1.21	1.26	0.90	80	7.89	20.05	2.34
Denmark	697	4.92	4.82	12.45	0.29	0.12	0.67	0.74	1.37	1.22	0.69	75	4.88	32.33	5.67
Hong Kong	2,657	5.08	10	12.15	0.26	0.33	0.44	0.88	1.67	1.45	0.95	73	11.25	22.96	12.62
India	885	2.81	2.42	11.42	0.38	0.20	0.59	0.86	1.13	1.04	0.25	61	3.52	0.43	0.67
Ireland	316	5.19	6.22	12.39	0.25	0.24	0.64	0.86	1.54	1.67	0.90	81	4.11	22.81	12.07
Italy	59	4.71	6.22	12.72	0.37	0.14	0.55	0.68	1.03	1.32	0.98	66	13.50	19.42	0.71
Malaysia	2,199	3.57	3.61	11.69	0.30	0.32	0.62	0.78	2.79	1.50	0.85	79	−2.04	4.06	3.54
New Zealand	330	4.96	10	12.09	0.24	0.12	0.47	0.90	1.73	1.50	0.96	80	13.63	15.47	3.58
Norway	418	4.94	6.22	12.71	0.20	0.22	0.60	0.84	1.71	1.33	0.94	75	10.85	35.59	2.36
Pakistan	55	1.44	1	10.09	0.41	0.40	0.75	0.78	1.07	1.00	0.18	73	−3.64	0.40	1.00
Singapore	1,556	4.71	4.82	11.88	0.29	0.28	0.47	0.88	2.55	2.06	0.91	79	13.88	23.60	9.37
South Africa	626	5.55	4.82	11.47	0.34	0.20	0.50	0.50	1.26	1.11	0.91	79	13.66	2.94	1.68
Sweden	370	5.86	4.82	12.53	0.29	0.37	0.51	0.89	1.09	1.14	0.88	83	7.84	25.43	8.29
U.K.	6,756	5.43	10	12.10	0.27	0.24	0.60	0.84	1.56	1.55	0.87	85	11.89	22.66	4.01
Mean (Total)	(17,837)	4.60	6.07	11.95	0.29	0.25	0.56	0.81	1.55	1.37	0.80	76.36	9.16	17.73	4.85
Mean (Total)—Strong home regime	(11,449)	5.07	8.38	12.26	0.25	0.22	0.54	0.83	1.49	1.44	0.93	77.14	11.23	22.71	5.38
Mean (Total)—Weak home regime	(6,388)	4.12	3.76	11.65	0.33	0.27	0.59	0.78	1.61	1.30	0.67	75.57	5.46	12.74	4.32
t-value for mean difference		43.70	390.00	10.05	−14.57	−0.26	−8.62	11.26	−25.13	0.89	24.37	55.16	87.51	89.12	16.00

(*continued on next page*)

TABLE 3 (continued)

The sample consists of 17,837 firm-year observations from 14 countries.

Variable Definitions:

AUDFEE = natural log of audit fee in thousands of U.S. dollars;
Wingate Index = Wingate's litigation index;
CROSS = 1 when a firm is cross-listed in a foreign country, and 0 otherwise;
LNTA = natural log of year-end total assets in thousands of U.S. dollars;
INVREC = sum of inventories and receivables divided by total assets;
LOSS = 1 when the firm reports a net loss, and 0 otherwise;
LEV = ratio of year-end total liabilities to total assets;
ISSUE = 1 when a firm has obtained long-term debt or equity financing by issuing bonds or new shares to outside capital suppliers in the past three years, and 0 otherwise;
NBS = natural log of 1 plus the number of business segments;
NGS = natural log of 1 plus the number of geographical segments;
BIG4 = 1 when a firm uses one of the Big 4 (5, 6, or 8) auditors, and 0 otherwise;
Raw DISCL = a country's disclosure level measured by the CIFAR index (not adjusted for the median value of the index);
B4DO = Big 4 dominance in the country, measured by the difference between the market share of the smallest Big 4 auditor and that of the largest non-Big 4 auditor;
GDP = gross domestic product per capita in thousands of U.S. dollars; and
FDI = foreign direct investment scaled by total GDP.

lower current assets ratio (*INVREC*), have less diversified business operations (*NBS*), have lower leverage (*LEV*), and engage more frequently in equity or bond issuance (*ISSUE*), compared with the weak-home-regime sample. Third, there is no significant difference in *LOSS* and *NGS* between the two samples. Last, the strong-home-regime sample has a higher disclosure level (*DISCL*), a greater degree of Big 4 dominance (*B4DO*), wealthier economies (*GDP*), and more foreign direct investment (*FDI*), compared with the weak-home-regime sample.

Table 4 presents the Pearson correlation coefficients and their p-values (in parentheses) for the variables in Equation (10). The magnitudes of the pairwise correlations among the firm-specific variables are no greater than 0.4, with the highest correlation between *NGS* and *NBS*. In contrast, *REGIME* is highly correlated with the four country-level control variables, with the highest correlation, 0.687, between *REGIME* and *B4DO*. Because of this high correlation, we also perform regression analyses without the four country-level control variables.

V. REGRESSION RESULTS

Existence of Cross-Listing Audit Fee Premium (H1 and H2)

Table 5 reports the ordinary least squares (OLS) estimates of the coefficients for Equation (10), with t-values in parentheses. The t-values are computed using robust standard errors adjusted for heteroscedasticity and clustering at the firm level (to control for serial correlations). Unless otherwise specified, all of the regressions are estimated after removing outliers that have a Cook's (1977) distance value greater than 4/(sample size). As a result, the actual sample size is slightly smaller than 17,837 and varies across the regressions. As a benchmark, we first estimate Equation (10) without *CROSS* using the full sample, and report the results in Columns (1a, 1b). We then estimate Equation (10) with *CROSS* included, using the full sample and various subsamples. The regressions in Columns (1a)–(6a) exclude the four country-level controls, whereas those in Columns (1b)–(4b) include them. The regressions in Columns (5a, 6a) include country indicator variables, but exclude the four country-level controls.

To test H1, we estimate Equation (10) for subsample H1 (i.e., the firms cross-listed in stronger foreign regimes and NCL firms in their home countries). Columns (2a, 5a, 2b) show that the coefficient on *CROSS* is highly significant with an expected positive sign ($t = 7.50$, $t = 8.33$, and $t = 10.33$, respectively), which supports H1. The audit fee premiums charged for the firms that are cross-listed in stronger foreign regimes are economically significant as well. For example, the results in Column (2b) translate into a cross-listing fee premium of 11.66 percent when all of the control variables are set at their mean values.

To test H2, we estimate Equation (10) for subsample H2 (i.e., the firms cross-listed in similar or weaker foreign regimes and NCL firms in their home countries). All of the firms from Denmark, India, Ireland, Italy, Pakistan, South Africa, and Sweden are excluded from this sample, because all of the CL firms from these countries are listed in stronger foreign regimes. Columns (3a, 6a, 3b) show that the coefficient on *CROSS* is statistically insignificant ($t = 1.33$, $t = 0.82$, and $t = 1.25$, respectively) and is much smaller in magnitude than that for subsample H1. Recall that H2 posits that the coefficient on *CROSS* for subsample H2 will be significantly positive if a cross-listing increases audit complexity. Therefore, the insignificant coefficient on *CROSS* in Columns (3a, 6a, 3b) suggests that cross-listings do not materially increase audit complexity.

Columns (2a, 5a, 2b) and (3a, 6a, 3b), taken as a whole, suggest that the observed cross-listing fee premiums are driven mainly by an upward shift in the strength of legal

TABLE 4
Correlations Matrix

Variable	*AUDFEE*	*REGIME*	*CROSS*	*LNTA*	*INVREC*	*LOSS*	*LEV*	*ISSUE*	*NBS*	*NGS*	*BIG4*	*DISCL*	*B4DO*	*GDP*	*FDI*
REGIME	0.400 (<0.001)	—													
CROSS	0.079 (<0.001)	−0.112 (<0.001)	—												
LNTA	0.470 (<0.001)	0.086 (<0.001)	0.088 (<0.001)	—											
INVREC	0.035 (<0.001)	−0.116 (<0.001)	−0.022 (0.004)	−0.219 (<0.001)	—										
LOSS	−0.066 (<0.001)	−0.002 (0.996)	0.001 (0.931)	−0.288 (<0.001)	−0.043 (<0.001)	—									
LEV	0.069 (<0.001)	−0.104 (<0.001)	0.012 (0.102)	−0.017 (0.025)	0.156 (<0.001)	0.122 (<0.001)	—								
ISSUE	0.158 (<0.001)	0.069 (<0.001)	0.016 (0.037)	0.189 (<0.001)	0.026 (<0.001)	0.052 (<0.001)	0.044 (<0.001)	—							
NBS	0.124 (<0.001)	−0.149 (<0.001)	0.124 (<0.001)	0.106 (<0.001)	0.019 (0.012)	0.019 (0.010)	0.013 (0.074)	0.072 (<0.001)	—						
NGS	0.218 (<0.001)	0.043 (<0.001)	0.066 (<0.001)	0.087 (<0.001)	0.009 (0.244)	0.028 (<0.001)	−0.000 (0.964)	0.088 (<0.001)	0.398 (<0.001)	—					
BIG4	0.248 (<0.001)	0.270 (<0.001)	0.009 (0.223)	0.200 (<0.001)	−0.054 (<0.001)	−0.071 (<0.001)	−0.061 (<0.001)	0.028 (<0.001)	0.040 (<0.001)	0.053 (<0.001)	—				
DISCL	0.305 (<0.001)	0.538 (<0.001)	−0.029 (<0.001)	0.044 (<0.001)	−0.073 (<0.001)	−0.004 (0.621)	−0.053 (<0.001)	−0.002 (0.765)	0.026 (0.001)	0.094 (<0.001)	0.249 (<0.001)	—			
B4DO	0.368 (<0.001)	0.687 (<0.001)	−0.167 (<0.001)	0.062 (<0.001)	−0.048 (<0.001)	−0.027 (<0.001)	−0.081 (<0.001)	0.040 (<0.001)	−0.159 (<0.001)	0.073 (<0.001)	0.174 (<0.001)	0.182 (<0.001)	—		
GDP	0.338 (<0.001)	0.656 (<0.001)	−0.103 (<0.001)	0.118 (<0.001)	−0.111 (<0.001)	−0.005 (0.517)	−0.065 (<0.001)	0.094 (<0.001)	−0.112 (<0.001)	0.079 (<0.001)	0.190 (<0.001)	−0.379 (<0.001)	0.117 (<0.001)	—	
FDI	0.101 (<0.001)	0.184 (<0.001)	0.011 (0.150)	0.073 (<0.001)	−0.033 (<0.001)	0.018 (0.019)	−0.054 (<0.001)	0.057 (<0.001)	−0.022 (0.003)	0.040 (<0.001)	0.126 (<0.001)	0.175 (<0.001)	0.198 (<0.001)	0.263 (<0.001)	—

(*continued on next page*)

TABLE 4 (continued)

The sample consists of 17,837 firm-year observations from 14 countries.

Variable Definitions:

AUDFEE = natural log of audit fee in thousands of U.S. dollars;
REGIME = natural log of Wingate's litigation index;
CROSS = 1 when a firm is cross-listed in a foreign country, and 0 otherwise;
LNTA = natural log of year-end total assets in thousands of U.S. dollars;
INVREC = sum of inventories and receivables divided by total assets;
LOSS = 1 when the firm reports a net loss, and 0 otherwise;
LEV = ratio of year-end total liabilities to total assets;
ISSUE = 1 when a firm has obtained long-term debt or equity financing by issuing bonds or new shares to outside capital suppliers in the past three years, and 0 otherwise;
NBS = natural log of 1 plus the number of business segments;
NGS = natural log of 1 plus the number of geographical segments;
BIG4 = 1 when a firm uses one of the Big 4 (5, 6, or 8) auditors, and 0 otherwise;
DISCL = a country's disclosure level measured by the (median-adjusted) CIFAR index;
B4DO = Big 4 dominance in the country, measured by the difference between the market share of the smallest Big 4 auditor and that of the largest non-Big 4 auditor;
GDP = gross domestic product per capita in thousands of U.S. dollars; and
FDI = foreign direct investment scaled by total GDP.

TABLE 5
Results of Multivariate Regressions for Testing H1 and H2

$$AUDFEE_{ijt} = \beta_0 + \beta_1 REGIME_j + \beta_2 CROSS_{ijt} + \delta_1 LNTA_{ijt} + \delta_2 INVREC_{ijt} + \delta_3 LOSS_{ijt} + \delta_4 LEV_{ijt} + \delta_5 ISSUE_{ijt} + \delta_6 NBS_{ijt} + \delta_7 NGS_{ijt} + \delta_8 BIG4_{ijt} + \varphi_1 DISCL_j + \varphi_2 B4DO_{jt} + \varphi_3 GDP_{jt} + \varphi_4 FDI_{jt} + \text{Fixed Effects} + \text{Error Term}$$

Panel A: Test Variables

	(1a)	(2a)	(3a)	(4a)	(5a)	(6a)	(1b)	(2b)	(3b)	4b
Variable (Pred. Sign)	Full Sample	Cross-Listed in Stronger Regimes	Cross-Listed in Weaker or Similar Regimes	Full Sample	Cross-Listed in Stronger Regimes	Cross-Listed in Weaker or Similar Regimes	Full Sample	Cross-Listed in Stronger Regimes	Cross-Listed in Weaker or Similar Regimes	Full Sample
REGIME (+)	1.357	1.391	1.636	1.391	1.419	1.962	0.772	0.751	1.209	0.758
	(41.42***)	(40.73***)	(38.79***)	(42.31***)	(43.85***)	(39.49***)	(17.09***)	(16.17***)	(23.84***)	(16.78***)
CROSS (+)		0.344	0.149	0.341	0.421	0.088		0.482	0.148	0.449
		(7.50***)	(1.33)	(7.75***)	(8.33***)	(0.82)		(10.33***)	(1.25)	(10.07***)

Panel B: Firm-Specific Control Variables (*FSCONTROL*)

	(1a)	(2a)	(3a)	(4a)	(5a)	(6a)	(1b)	(2b)	(3b)	4b
Variable (Pred. Sign)	Full Sample	Cross-Listed in Stronger Regimes	Cross-Listed in Weaker or Similar Regimes	Full Sample	Cross-Listed in Stronger Regimes	Cross-Listed in Weaker or Similar Regimes	Full Sample	Cross-Listed in Stronger Regimes	Cross-Listed in Weaker or Similar Regimes	Full Sample
LNTA (+)	0.418	0.402	0.397	0.409	0.405	0.416	0.423	0.410	0.405	0.417
	(43.89***)	(40.58***)	(34.37***)	(43.40***)	(41.24***)	(38.02***)	(43.39***)	(41.20***)	(34.63***)	(43.73***)
INVREC (+)	1.257	1.239	1.272	1.264	1.276	1.353	1.298	1.285	1.284	1.316
	(18.87***)	(17.95***)	(16.78***)	(19.00***)	(19.08***)	(18.03***)	(19.13***)	(18.62***)	(16.74***)	(19.45***)
LOSS (+)	0.206	0.185	0.199	0.196	0.227	0.223	0.216	0.213	0.235	0.219
	(7.91***)	(6.84***)	(6.71***)	(7.50***)	(8.75***)	(7.97***)	(8.41***)	(8.14***)	(8.04***)	(8.54***)
LEV (+)	0.705	0.696	0.731	0.704	0.589	0.717	0.711	0.685	0.723	0.699
	(15.86***)	(15.68***)	(14.49***)	(15.90***)	(14.92***)	(14.89***)	(16.44***)	(15.60***)	(14.47***)	(16.15***)
ISSUE (+)	0.145	0.144	0.261	0.140	0.188	0.268	0.142	0.148	0.255	0.143
	(4.32***)	(4.18***)	(6.90***)	(4.21***)	(5.62***)	(7.31***)	(4.24***)	(4.31***)	(6.67***)	(4.32***)
NBS (+)	0.081	0.078	0.122	0.078	0.137	0.126	0.088	0.080	0.127	0.079
	(8.80***)	(8.18***)	(11.28***)	(8.47***)	(13.86***)	(12.34***)	(9.88***)	(8.74***)	(12.11***)	(8.98***)

(*continued on next page*)

TABLE 5 (continued)

Variable (Pred. Sign)	(1a) Full Sample	(2a) Cross-Listed in Stronger Regimes	(3a) Cross-Listed in Weaker or Similar Regimes	(4a) Full Sample	(5a) Cross-Listed in Stronger Regimes	(6a) Cross-Listed in Weaker or Similar Regimes	(1b) Full Sample	(2b) Cross-Listed in Stronger Regimes	(3b) Cross-Listed in Weaker or Similar Regimes	4b Full Sample
NGS (+)	0.219 (20.08***)	0.218 (19.17***)	0.234 (18.83***)	0.219 (20.02***)	0.199 (16.38***)	0.201 (16.17***)	0.184 (16.25***)	0.178 (15.28***)	0.201 (15.78***)	0.177 (15.72***)
BIG4 (+)	0.273 (6.98***)	0.282 (6.93***)	0.137 (3.12***)	0.270 (6.83***)	0.207 (5.40***)	0.135 (3.21***)	0.215 (5.74***)	0.211 (5.33***)	0.157 (3.49***)	0.217 (5.67***)

Panel C: Country-Level Control Variables (*CLCONTROL*)

Variable (Pred. Sign)	(1a) Full Sample	(2a) Cross-Listed in Stronger Regimes	(3a) Cross-Listed in Weaker or Similar Regimes	(4a) Full Sample	(5a) Cross-Listed in Stronger Regimes	(6a) Cross-Listed in Weaker or Similar Regimes	(1b) Full Sample	(2b) Cross-Listed in Stronger Regimes	(3b) Cross-Listed in Weaker or Similar Regimes	4b Full Sample
DISCL (+)							0.030 (10.79***)	0.029 (10.14***)	0.028 (7.42***)	0.029 (10.65***)
B4DO (+)							0.044 (9.93***)	0.051 (11.20***)	0.033 (5.72***)	0.050 (11.45***)
GDP (+)							0.008 (3.43***)	0.007 (3.26***)	0.006 (1.88*)	0.008 (3.47***)
FDI (+)							0.003 (2.15**)	0.002 (1.31)	0.007 (5.02***)	0.002 (1.34)
Intercept	−4.378 (−33.46***)	−4.270 (−31.27***)	−4.890 (−30.63***)	−4.353 (−33.40***)	−4.328 (−33.71***)	−5.786 (−35.07***)	−3.819 (−28.95***)	−3.677 (−27.18***)	−4.633 (−28.81***)	−3.794 (−29.20***)
Country Indicators	Not Included	Not Included	Not Included	Not Included	Included	Included	Not Included	Not Included	Not Included	Not Included
Adj. R^2	0.573	0.553	0.548	0.574	0.622	0.604	0.605	0.589	0.572	0.609
n	17,837	17,646	13,655	17,837	17,646	13,655	17,837	17,646	13,655	17,837

(*continued on next page*)

TABLE 5 (continued)

---

*, **, *** Indicates statistical significance at the 10 percent, 5 percent, and 1 percent levels (two-tailed), respectively.

The sample consists of 17,837 firm-year observations from 14 countries. Tests in Columns (2a), (5a), and (2b) are performed for subsample H1. For example, the sample of CL firms that are listed in foreign countries whose legal regimes are stronger than those of the home countries and NCL firms in the home countries. The tests in Columns (3a), (6a), and (3b) are performed for subsample H2. For example, the sample of CL firms that are listed in foreign countries whose legal regimes are no stronger than those of the home countries and NCL firms in the home countries.

Variable Definitions:
AUDFEE = natural log of audit fee in thousands of U.S. dollars;
REGIME = natural log of Wingate's litigation index;
CROSS = 1 when a firm is cross-listed in a foreign country, and 0 otherwise;
LNTA = natural log of year-end total assets in thousands of U.S. dollars;
INVREC = sum of inventories and receivables divided by total assets;
LOSS = 1 when the firm reports a net loss, and 0 otherwise;
LEV = ratio of year-end total liabilities to total assets;
ISSUE = 1 when a firm has obtained long-term debt or equity financing by issuing bonds or new shares to outside capital suppliers in the past three years, and 0 otherwise;
NBS = natural log of 1 plus the number of business segments;
NGS = natural log of 1 plus the number of geographical segments;
BIG4 = 1 when a firm uses one of the Big 4 (5, 6, or 8) auditors, and 0 otherwise;
DISCL = a country's disclosure level measured by the (median-adjusted) CIFAR index;
B4DO = Big 4 dominance in the country, measured by the difference between the market share of the smallest Big 4 auditor and that of the largest non-Big 4 auditor;
GDP = gross domestic product per capita in thousands of U.S. dollars; and
FDI = foreign direct investment scaled by total GDP.

regimes, but not by a change in audit complexity.[14] The significant coefficient on *CROSS* in the full sample regressions in Columns (4a, 4b) is caused by the cross-listings in countries with stronger legal regimes.

With respect to the coefficients on the other variables, the following are noteworthy. First, the coefficient on *REGIME* is highly significant with an expected positive sign in all columns. This indicates that the strength of the home country's legal regime is an important determinant of the audit fee, a finding consistent with Choi et al. (2008). Second, audit fees increase with client size (*LNTA*), client complexity (*INVREC*, *NBS*, and *NGS*), client-specific risk (*LEV* and *LOSS*), financial transactions (*ISSUE*), and auditor quality (*BIG4*). These findings are consistent with previous audit fee studies (e.g., Simunic 1980; Seetharaman et al. 2002). Third, as shown in Columns (1b)–(4b), the coefficients on the country-level control variables are significant in most cases, which suggests that international audit fee research such as ours should control for country-wide, macroeconomic variables.

Determinant of the Magnitude of the Cross-Listing Audit Fee Premium (H3)

We find significant audit fee premiums only when CL firms submit themselves to stronger foreign legal regimes. For such CL firms, H3 predicts that the cross-listing fee premiums increase with *the difference* between the strength of the legal regime of the cross-listed foreign country and that of the home country. As explained before, we test H3 using two data sets, separately, in which the strength of the (cross-listed) foreign legal regime is fixed at two different benchmarks, and report the regression results in Table 6. Specifically, Columns (1a, 3a, 1b) report the results for the CL firms that are cross-listed in the U.S. and NCL firms in their respective home countries. Columns (2a, 4a, 2b) report the results for the CL firms that are cross-listed in the four countries with *REGIME* = 2.30 but not cross-listed in the U.S. at the same time, and NCL firms in their respective home countries. The regressions in Columns (1a)–(4a) do not include the four country-level controls, whereas those in Columns (1b, 2b) do. The regressions in Columns (3a, 4a) include country indicator variables, whereas those in the other columns do not.

As shown in Table 6, across all columns, the coefficient on *CROSS* is significant with an expected positive sign, and the coefficient on the interaction term, *CROSS * REGIME*, is significant with an expected negative sign. Consistent with H3, these results indicate that the cross-listing fee premiums for firms cross-listed in stronger foreign regimes are larger (smaller) when those firms come from countries with relatively weak (strong) legal regimes.

To gain further insights, note from Columns (1b, 2b) that an average firm in Pakistan (with a Wingate index of 1 and other variables set at their mean values) pays a cross-listing fee premium of about 33 percent when it is cross-listed in the U.S. This premium would be about 23 percent if the same firm were cross-listed in the U.K. (with a Wingate index of 10). Similarly, an average firm in the countries with the Wingate index being 4.82 (i.e., Denmark, Singapore, South Africa, and Sweden) pays a premium of about 6.9 percent when it is cross-listed in the U.S., whereas this premium falls to about 3.54 percent when it is cross-listed in the U.K. In short, our results suggest that the cross-listing fee premium is economically significant and is positively related to the difference in the strength of legal regimes between the cross-listed foreign country and home country.

---

[14] This is consistent with the argument in Seetharaman et al. (2002) that the legal liability is the main reason for the increase in audit fee that is associated with U.K. firms being cross-listed in the U.S. There is a possibility that audit complexity increases *only when* a firm is cross-listed in stronger foreign regimes, not when a firm is cross-listed in weaker foreign regimes, in which case we may observe the same phenomena as documented in Table 5.

TABLE 6
Results of Multivariate Regressions for Testing H3

$$AUDFEE_{ijt} = \beta_0 + \beta_1 REGIME_j + \beta_2 CROSS_{ijt} + \beta_3(CROSS_{ijt} * REGIME_j) + \delta_1 LNTA_{ijt} + \delta_2 INVREC_{ijt} + \delta_3 LOSS_{ijt} + \delta_4 LEV_{ijt} + \delta_5 ISSUE_{ijt} + \delta_6 NBS_{ijt} + \delta_7 NGS_{ijt} + \delta_8 BIG4_{ijt} + \varphi_1 DISCL_j + \varphi_2 B4DO_{jt} + \varphi_3 GDP_{jt} + \varphi_4 FDI_{jt} + \text{Fixed Effects} + \text{Error Term}$$

Panel A: Test Variables

	(1a)	(2a)	(3a)	(4a)	(1b)	(2b)
Variable (Pred. Sign)	**Cross-Listed in the U.S.**	**Cross-Listed in Countries where *REGIME* = 2.30**	**Cross-Listed in the U.S.**	**Cross-Listed in Countries where *REGIME* = 2.30**	**Cross-Listed in the U.S.**	**Cross-Listed in Countries where *REGIME* = 2.30**
REGIME (+)	1.394	3.957	1.582	1.863	0.972	3.807
	(36.21***)	(30.05***)	(42.25***)	(6.19***)	(14.39***)	(12.27***)
CROSS (+)	0.822	3.698	0.965	1.834	0.829	2.857
	(4.77***)	(5.58***)	(6.15***)	(3.57***)	(6.56***)	(4.96***)
CROSS * *REGIME* (−)	−0.362	−2.164	−0.326	−0.881	−0.343	−1.597
	(−3.83***)	(−5.58***)	(−3.80***)	(−2.78***)	(−3.81***)	(−4.44***)

Panel B: Firm-Specific Control Variables (*FSCONTROL*)

	(1a)	(2a)	(3a)	(4a)	(1b)	(2b)
Variable (Pred. Sign)	**Cross-Listed in the U.S.**	**Cross-Listed in Countries where *REGIME* = 2.30**	**Cross-Listed in the U.S.**	**Cross-Listed in Countries where *REGIME* = 2.30**	**Cross-Listed in the U.S.**	**Cross-Listed in Countries where *REGIME* = 2.30**
LNTA (+)	0.399	0.280	0.400	0.273	0.408	0.276
	(37.93***)	(21.34***)	(38.80***)	(20.88***)	(38.99***)	(20.64***)
INVREC (+)	1.253	0.831	1.258	0.753	1.294	0.745
	(17.49***)	(9.11***)	(18.38***)	(8.49***)	(18.16***)	(8.62***)
LOSS (+)	0.204	0.096	0.232	0.091	0.226	0.018
	(7.22***)	(2.89***)	(8.81***)	(2.79***)	(8.37***)	(0.56)
LEV (+)	0.700	0.132	0.657	0.153	0.687	0.257
	(15.06***)	(3.30***)	(14.60***)	(3.88***)	(14.80***)	(6.79***)
ISSUE (+)	0.145	−0.071	0.216	−0.070	0.163	−0.092
	(3.97***)	(−1.76*)	(6.26***)	(−1.75*)	(4.57***)	(−2.30**)

(continued on next page)

TABLE 6 (continued)

Variable (Pred. Sign)	(1a) Cross-Listed in the U.S.	(2a) Cross-Listed in Countries where *REGIME* = 2.30	(3a) Cross-Listed in the U.S.	(4a) Cross-Listed in Countries where *REGIME* = 2.30	(1b) Cross-Listed in the U.S.	(2b) Cross-Listed in Countries where *REGIME* = 2.30
NBS (+)	0.080 (7.87***)	0.095 (7.81***)	0.113 (11.42***)	0.118 (9.26***)	0.074 (7.65***)	0.100 (7.97***)
NGS (+)	0.210 (17.96***)	0.148 (8.79***)	0.186 (15.40***)	0.140 (7.66***)	0.169 (14.10***)	0.135 (8.03***)
BIG4 (+)	0.277 (6.50***)	0.036 (0.73)	0.215 (5.17***)	0.064 (1.33)	0.192 (4.65***)	−0.059 (−1.24)

Panel C: Country-Level Control Variables (*CLCONTROL*)

Variable (Pred. Sign)	(1a) Cross-Listed in the U.S.	(2a) Cross-Listed in Countries where *REGIME* = 2.30	(3a) Cross-Listed in the U.S.	(4a) Cross-Listed in Countries where *REGIME* = 2.30	(1b) Cross-Listed in the U.S.	(2b) Cross-Listed in Countries where *REGIME* = 2.30
DISCL (+)					0.030 (10.08***)	0.079 (7.13***)
B4DO (+)					0.051 (10.50***)	0.029 (5.46***)
GDP (+)					0.008 (3.39***)	−0.003 (−1.03)
FDI (+)					0.003 (1.76*)	−0.034 (−12.01***)
Intercept	−4.227 (−28.60***)	−5.637 (−23.79***)	−4.772 (−27.07***)	−2.278 (−4.56***)	−3.606 (−24.86***)	−5.123 (−12.15***)
Country indicators	Not Included	Not Included	Included	Included	Not Included	Not Included
Adj. R^2	0.5323	0.582	0.595	0.615	0.571	0.620
n	16,851	5,495	16,851	5,495	16,851	5,495

(continued on next page)

TABLE 6 (continued)

*, **, *** Indicate statistical significance at the 10 percent, 5 percent, and 1 percent levels (two-tailed), respectively.

The tests in Columns (1a), (3a), and (1b) are performed with the reduced sample of firms that are cross-listed in the U.S. and NCL firms in their home countries. The tests in Columns (2a), (4a), and (2b) are performed with the reduced sample of firms that are cross-listed in countries in which *REGIME* is 2.30 or equivalently the Wingate index is 10 (including Australia, Hong Kong, New Zealand, and the U.K.) and NCL firms in their home countries.

Variable Definitions:

AUDFEE = natural log of audit fee in thousands of U.S. dollars;
REGIME = natural log of Wingate's litigation index;
CROSS = 1 when a firm is cross-listed in a foreign country, and 0 otherwise;
LNTA = natural log of year-end total assets in thousands of U.S. dollars;
INVREC = sum of inventories and receivables divided by total assets;
LOSS = 1 when the firm reports a net loss, and 0 otherwise;
LEV = ratio of year-end total liabilities to total assets;
ISSUE = 1 when a firm has obtained long-term debt or equity financing by issuing bonds or new shares to outside capital suppliers in the past three years, and 0 otherwise;
NBS = natural log of 1 plus the number of business segments;
NGS = natural log of 1 plus the number of geographical segments;
BIG4 = 1 when a firm uses one of the Big 4 (5, 6, or 8) auditors, and 0 otherwise;
DISCL = a country's disclosure level measured by the (median-adjusted) CIFAR index;
B4DO = Big 4 dominance in the country, measured by the difference between the market share of the smallest Big 4 auditor and that of the largest non-Big 4 auditor;
GDP = gross domestic product per capita in thousands of U.S. dollars; and
FDI = foreign direct investment scaled by total GDP.

Test of H3 Using Partitioned Samples

The regressions in Table 6 implicitly assume that the relations between the audit fee and its determinants other than *CROSS* do not differ significantly between the countries with weak and strong regimes. However, our univariate tests for the mean differences between the weak- and the strong-home-regime samples, as reported in Table 3, show that all of the explanatory variables in Equation (10), except *LOSS* and *NGS*, differ between the two samples. Therefore, the effects of these explanatory variables on the audit fee, which are captured by the regression coefficients, may differ systematically between the countries with weak and strong regimes.

To address this issue, we partition the sample of firms that are cross-listed in the U.S. and NCL firms (i.e., the sample used in Columns (1a, 3a, 1b) of Table 6) into two subsamples using the Wingate index of 6.22 as a cut-off value, namely: (1) a strong-home-regime sample (i.e., from the seven countries with the Wingate index greater than or equal to 6.22), and (2) a weak-home-regime sample (i.e., from the seven countries with the Wingate index less than 6.22).[15] We then estimate the following regression for each subsample.[16]

$$\begin{aligned} AUDFEE = {} & \beta_0 + \beta_1 CROSS + \sum_{k=1}^{8} \delta_k FSCONTROL + \sum_{k=1}^{4} \varphi_k CLCONTROL_k \\ & + \text{Fixed Effects} + \text{Error Term}. \end{aligned} \tag{12}$$

Columns (1) and (2) of Table 7 present the results for the weak- and strong-home-regime samples, and Column (3) reports the differences in the regression coefficients between the two samples. While the coefficient on *CROSS* (β_1) is significantly positive, it is much smaller for the strong-home-regime sample ($\beta_1 = 0.264$) than for the weak-home-regime sample ($\beta_1 = 0.470$). The difference in β_1 is significant at less than the 5 percent level ($t = -2.36$).[17] This means that the audit fee premiums associated with cross-listings in the U.S. are significantly smaller for CL firms *from* strong-home-regime countries than they are for those *from* weak-home-regime countries. This finding is consistent with the negative coefficient on *CROSS* * *REGIME* in Table 6, and lends further support to H3. We also repeat this analysis with the sample used in Columns (2a, 4a, 2b) of Table 6 and find similar results.

With respect to the coefficients on the firm-specific control variables, as shown in Table 7, they are larger for the strong-home-regime sample than they are for the weak-home-regime sample, except for *BIG4*. Consistent with Choi et al. (2008), we find that the Big 4 auditor fee premium is smaller in countries with stronger legal regimes than in those with weak legal regimes. In addition, we observe significant differences in the coefficients on

---

[15] The strong- and weak-home-regime samples are slightly different from those involved in the last three rows of Table 3 because here we consider only the firms that are cross-listed in the U.S. and corresponding NCL firms.

[16] We do not include *REGIME* in Equation (12) to mitigate potential multicollinearity problems, because *REGIME* is highly correlated with the country-level controls and has a relatively small (cross-country) variation within each subsample. The result on our test variable remains unchanged when *REGIME* is included, with the difference in the coefficient on *CROSS* being –0.223 (t = –2.09**).

[17] When we perform country-by-country regressions without the country-level controls, the results are qualitatively similar. For example, the average (across countries) coefficient on the *CROSS* variable is 0.528 for the weak-home-regime sample, while it is 0.268 for the strong-home-regime sample.

TABLE 7

Tests for Differences in Regression Coefficients between Strong- and Weak-Home-Regime Subsamples for the Firms Cross-Listed in the U.S. and Corresponding NCL Firms

$$\begin{aligned} AUDFEE_{ijt} = {} & \beta_0 + \beta_1 CROSS_{ijt} + \delta_1 LNTA_{ijt} + \delta_2 INVREC_{ijt} + \delta_3 LOSS_{ijt} + \delta_4 LEV_{ijt} \\ & + \delta_5 ISSUE_{ijt} + \delta_6 NBS_{ijt} + \delta_7 NGS_{ijt} + \delta_8 BIG4_{ijt} + \varphi_1 DISCL_j + \varphi_2 B4DO_{jt} \\ & + \varphi_3 GDP_{jt} + \varphi_4 FDI_{jt} + \text{Fixed Effects} + \text{Error Term} \end{aligned}$$

Panel A: Test Variables

	Sample Partition by the Wingate Litigation Index		
Variable (Pred. Sign)	**(1) Strong-Home-Regime Sample**	**(2) Weak-Home-Regime Sample**	**(3) Difference = 1 − 2**
CROSS (+)	0.264	0.470	−0.206
	(2.66***)	(3.07***)	(−2.36**)

Panel B: Firm-Specific Control Variables (*FSCONTROL*)

	Sample Partition by the Wingate Litigation Index		
Variable (Pred. Sign)	**(1) Strong-Home-Regime Sample**	**(2) Weak-Home-Regime Sample**	**(3) Difference = 1 − 2**
LNTA (+)	0.424	0.292	0.129
	(26.90***)	(14.21***)	(4.98***)
INVREC (+)	1.352	0.526	0.826
	(12.96***)	(3.66***)	(4.66***)
LOSS (+)	0.400	0.020	0.380
	(10.05***)	(0.40)	(6.01***)
LEV (+)	0.564	0.229	0.335
	(8.88***)	(3.68***)	(3.76***)
ISSUE (+)	0.307	−0.182	0.489
	(5.52***)	(−2.72***)	(5.62***)
NBS (+)	0.112	0.067	0.045
	(6.48***)	(3.52***)	(1.75*)
NGS (+)	0.191	0.090	0.101
	(10.42***)	(3.26***)	(3.06***)
BIG4 (+)	0.081	0.213	−0.132
	(1.03)	(3.01***)	(−1.24)

Panel C: Country-Level Control Variables (*CLCONTROL*)

	Sample Partition by the Wingate Litigation Index		
Variable (Pred. Sign)	**(1) Strong-Home-Regime Sample**	**(2) Weak-Home-Regime Sample**	**(3) Difference = 1 − 2**
DISCL (+)	0.044	0.062	−0.018
	(10.50***)	(10.12***)	(−2.39**)
B4DO (+)	−0.052	0.089	−0.147
	(−2.25**)	(13.48***)	(−5.89***)
GDP (+)	−0.023	0.022	−0.045
	(−4.41***)	(6.66***)	(−7.27***)
FDI (+)	0.010	−0.057	0.066
	(4.25***)	(−10.68***)	(11.50***)

(*continued on next page*)

TABLE 7 (continued)

Variable (Pred. Sign)	Sample Partition by the Wingate Litigation Index (1) Strong-Home-Regime Sample	(2) Weak-Home-Regime Sample	(3) Difference = 1 − 2
Intercept	−0.490 (−1.35)	−0.109 (−0.43)	−0.381 (−0.86)
Adj. R^2	0.412	0.441	
n	11,052	5,799	

*, **, *** Indicate statistical significance at the 10 percent, 5 percent, and 1 percent levels (two-tailed), respectively.

The sample consists of 16,851 firm-year observations from 14 countries. The strong-home-regime sample includes seven countries with the Wingate index being greater than or equal to 6.22 (i.e., Australia, Hong Kong, Ireland, Italy, New Zealand, Norway, and the U.K.). The weak-home-regime sample includes seven countries with the Wingate index being less than 6.22 (i.e., Denmark, India, Malaysia, Pakistan, Singapore, South Africa, and Sweden).

Variable Definitions:
AUDFEE = natural log of audit fee in thousands of U.S. dollars;
REGIME = natural log of Wingate's litigation index;
CROSS = 1 when a firm is cross-listed in a foreign country, and 0 otherwise;
LNTA = natural log of year-end total assets in thousands of U.S. dollars;
INVREC = sum of inventories and receivables divided by total assets;
LOSS = 1 when the firm reports a net loss, and 0 otherwise;
LEV = ratio of year-end total liabilities to total assets;
ISSUE = 1 when a firm has obtained long-term debt or equity financing by issuing bonds or new shares to outside capital suppliers in the past three years, and 0 otherwise;
NBS = natural log of 1 plus the number of business segments;
NGS = natural log of 1 plus the number of geographical segments;
BIG4 = 1 when a firm uses one of the Big 4 (5, 6, or 8) auditors, and 0 otherwise;
DISCL = a country's disclosure level measured by the (median-adjusted) CIFAR index;
B4DO = Big 4 dominance in the country, measured by the difference between the market share of the smallest Big 4 auditor and that of the largest non-Big 4 auditor;
GDP = gross domestic product per capita in thousands of U.S. dollars; and
FDI = foreign direct investment scaled by total GDP.

the four country-level controls between the two samples.[18] Overall, we find that the audit fee structure differs between countries with weak and strong legal regimes.

VI. SENSITIVITY ANALYSES

We perform a variety of sensitivity analyses to check whether our regression results are robust to alternative variable definitions, econometric methods, and model specifications.

---

[18] Note that *ISSUE* has an unexpected negative coefficient for the weak-home-regime sample. To the extent that audit fees reflect the demand for quality audit services, this evidence suggests that issuing firms have less demand for quality audits in countries with weak legal regimes. Among the country-level variables, *B4DO* and *GDP* have unexpected negative coefficients for the strong-home-regime sample. The negative coefficient on *B4DO* suggests that in countries with strong legal regimes, the high concentration of Big 4 auditors may lead to greater competition among these auditors, and thus lower audit fees. The negative coefficient on *GDP* suggests a negative association between auditor remuneration and the wealth level of the countries that have strong legal regimes. In addition, *FDI* has an unexpected negative sign for the weak-home-regime sample. This suggests that in the countries with weak legal environments, auditors may have more incentives to compete for audit engagements associated with foreign investments, and thus charge lower fees when the market is more open to foreign investments.

Cross-Listings on Major Stock Exchanges versus Other Foreign Trading

Although cross-listed firms must have foreign-equity trading, a firm's shares can be traded in foreign countries without a cross-listing on their major stock exchanges, for example, through private placements of equity shares (e.g., Lang, Raedy, and Yetman 2003, 379–380). In this study, we treat both the cross-listed firms and the firms whose equities are traded in other countries, but without cross-listings, as "cross-listed firms," due to the difficulty in separating these two cases. Therefore, our results may include some noise. It should be pointed out, however, that this noise introduces a conservative bias against rejecting the null hypothesis of no cross-listing effect. To understand this, note that for foreign firms that are traded in a country but not cross-listed on this country's major exchanges, the auditor's legal liability that is associated with financial statement audits is less affected by this country's legal regime, compared with a situation in which those foreign firms are cross-listed on this country's major exchanges. Consistent with this argument, Lang, Raedy, and Yetman (2003) find that various characteristics of the non-U.S. firms that are traded in the U.S. OTC market are significantly different from those that are traded on the NYSE, AMEX, and NASDAQ. They explain that trading in the OTC market provides a setting in which foreign firms partially enjoy the benefit of cross-listings without the full exposure to the legal requirements that are imposed on the CL firms traded on the major exchanges.

Although we do not have data that distinguish cross-listings and the other foreign trading in the non-U.S. countries, we are able to separate for our sample the foreign firms that are cross-listed on the major U.S. exchanges (NYSE, AMEX, and NASDAQ) from those that are traded in the other U.S. markets such as the OTC market. We construct a reduced sample of 14,941 (CL and NCL) firm-year observations from a set of non-U.S. countries, each of which has at least one firm being traded on the major U.S. exchanges. This sample excludes observations that are cross-listed in non-U.S. countries, and contains 162 observations that are cross-listed on the major U.S. exchanges and 486 observations that are traded in the U.S., but not listed on the major exchanges. We use this sample to estimate the following regression

$$\begin{aligned} AUDFEE = {} & \beta_0 + \beta_1 REGIME + \beta_2 CROSS + \beta_3 USEXCHANGE \\ & + \sum_{k=1}^{8} \delta_k FSCONTROL + \sum_{k=1}^{2} \varphi_k CLCONTROL_k \\ & + \text{Fixed Effects} + \text{Error Term}. \end{aligned} \tag{13}$$

USEXCHANGE is an indicator variable that equals 1 if the firm's shares are traded on the major U.S. exchanges, and 0 otherwise (i.e., if the shares are traded in the U.S., but not on the NYSE, AMEX, or NASDAQ, or if they are traded only in their home countries). As a result, the coefficient on *CROSS* (β_2) captures the audit fee premiums for foreign firms that are traded in the U.S., but not on the major exchanges, whereas the coefficient on *USEXCHANGE* (β_3) captures the incremental fee premiums paid by foreign firms traded on the major U.S. exchanges. The sum of β_2 and β_3 captures the total cross-listing fee premiums paid by the foreign firms that are traded on the major U.S. exchanges.

Though not tabulated, we find that the coefficients (t-values) on *REGIME*, *CROSS*, and *USEXCHANGE* are $\beta_1 = 0.693$ (15.51), $\beta_2 = 0.190$ (3.83), and $\beta_3 = 0.441$ (4.66), respectively. The significantly positive coefficients on *CROSS* and *USEXCHANGE* confirm our earlier finding that any foreign firms traded in the U.S., whether on the major exchanges or not, pay audit fee premiums, compared with the firms traded only in their home markets.[19] In addition, the significantly positive β_3 indicates that foreign firms whose shares are traded on the major U.S. exchanges pay even higher audit fee premiums, compared with the foreign firms that are traded in the U.S. but not on the major exchanges. This supports the view that our analyses that do not distinguish the two types of foreign trading introduce a conservative bias against rejecting the null hypothesis of no cross-listing effect.

Changes in Audit Fees Before and After Cross-Listings

The Worldscope database shows only the *current status* of cross-listings; thus, we do not know when firms started cross-listing their shares in foreign markets. In our empirical tests, the *CROSS* variable is coded 1 if the firm's shares were traded in foreign countries in the last year of our sample period (i.e., 2002), and 0 otherwise. As a result, we do not have a complete panel of cross-listings, which may cause a measurement error in our regressions. To mitigate this problem, we perform error-in-variable regressions (Greene 1997), and find that the results are qualitatively the same. To further alleviate the concern over incomplete panel data, we investigate how the audit fees change before and after cross-listings using a reduced sample of firm-years for which we can identify the exact cross-listing dates.

By searching the Citibank ADR list and other sources, we identify 102 non-U.S. companies that have the exact starting dates for cross-listing in the U.S. and are included in the initial data set we retrieved from Worldscope.[20] However, only 29 of these companies *began* their cross-listings during our sample period. The sample is further reduced to 11 companies due to unavailability of data on audit fee and other necessary financial information. For the 11 companies from five countries (Hong Kong, India, Singapore, South Africa, and the U.K.), we identify 27 firm-year observations before cross-listings and 26 observations after.

The average *AUDFEE* of these observations is 7.321 before cross-listings and 7.882 after. We compare this *AUDFEE* with the average *AUDFEE* for the NCL firms matched by country, year, and industry. The comparison reveals that a CL firm pays a higher audit fee (measured by *AUDFEE*) by an amount of 1.012 than a NCL firm before a cross-listing, and that the fee premium increases to 1.677 after the cross-listing. Although the change in the audit fee premium is in the expected direction, it is only marginally significant ($t = 1.50$, $p = 0.07$ in a one-tailed test), possibly due to the small sample size of only 53.[21]

---

[19] Our results are consistent with the findings of prior studies. Seetharaman et al. (2002) show that the audit fees paid by the U.K. firms traded on the U.S. OTC market are significantly higher than those of NCL firms. Although Lang, Raedy, and Yetman (2003) do not perform statistical tests, their Tables 3 and 6 suggest that there are differences between NCL firms and the OTC-traded foreign firms.

[20] The 527 firm-years associated with the 102 companies account for 54 percent of our sample of 980 firm-years cross-listed in the U.S. (see Table 2). We are unable to identify the exact cross-listing dates for other foreign companies in the U.S. due in large part to the difference in company names used in our data set versus ADR list.

[21] We also tried to perform similar tests with the data of cross-listings in the non-U.S. countries that were collected by Sarkissian and Schill (2004). Out of 107 of their firms that began cross-listing during our sample period, 24 are matched with the initial data we retrieved from Worldscope. After imposing necessary data requirements, we have two firms that are cross-listed in a stronger non-U.S. foreign legal regime and four firms that are cross-listed in similar or weaker foreign legal regimes. Thus, we are not able to perform a meaningful analysis using these few data.

Interactions between *CROSS* and Control Variables

When estimating Equation (11), we assume that *CROSS* interacts only with *REGIME.* It is possible, however, that *CROSS* interacts with some of the control variables in determining audit fees. To examine this possibility, we estimate Equation (14) using the same two samples as we used for testing H3:

$$\begin{aligned} AUDFEE = {} & \beta_0 + \beta_1 REGIME + \beta_2 CROSS + \beta_3 CROSS * REGIME \\ & + \sum_{k=1}^{8} \delta_k FSCONTROL + \sum_{k=1}^{4} \varphi_k CLCONTROL_k \\ & + \sum_{k=1}^{8} \theta_k CROSS * FSCONTROL \qquad (14) \\ & + \sum_{k=1}^{4} \xi_k CROSS * CLCONTROL_k + \text{Fixed Effects} \\ & + \text{Error Term.} \end{aligned}$$

Though not tabulated for brevity, we find that the coefficient on *CROSS * REGIME* remains significant with a negative sign, despite a few significant interactions between *CROSS* and control variables. For example, when we use the first sample (i.e., the firms that are cross-listed in the U.S. and NCL firms in their home countries), the significant interactions exist only for *CROSS * REGIME* (−0.421 with t = −3.48), *CROSS * LOSS* (−0.160 with t = −2.45), *CROSS * BIG4* (0.396 with t = 4.35), and *CROSS * DISCL* (0.031 with t = 2.32). None of the other interactions is significant. In short, our result with respect to H3 (i.e., the negative coefficient on *CROSS * REGIME*) is robust with respect to whether the interactions of *CROSS* with firm-level and country-level control variables are considered.[22]

Endogenous Cross-Listing Decision

A firm's cross-listing decision may not be exogenous. To address this potential self-selection problem, we apply the two-stage treatment effect approach. In the first stage, we estimate a probit model using *CROSS* as the dependent variable and the variables that are known to affect a firm's cross-listing decision as the explanatory variables (Doidge et al. 2004; Pagano et al. 2002), and then obtain the inverse Mills ratio. In the second stage, we estimate Equations (10) and (11) after adding the inverse Mills ratio as an additional control variable.

Specifically, we use the following model for the first-stage probit regression:

$$\begin{aligned} CROSS = {} & \beta_0 + \beta_1 REGIME + \beta_2 LNTA + \beta_3 SAG + \beta_4 NBS + \beta_5 NGS + \beta_6 BIG4 \\ & + \beta_7 GDP + \beta_8 DISCL + \text{Country Indicators} + \text{Error Term} \end{aligned}$$

where *SAG* denotes sales growth, which is calculated as the change in sales between years $t-1$ and t, divided by the sales in year $t-1$. The definitions of the other variables are the same as before (see Table 1). Our choice of explanatory variables is guided by Doidge et al. (2004) and Lang, Lins, and Miller (2003), except for *NBS*, *NGS*, and *BIG4*. We include

---

[22] We also find that adding the interaction variables, *CROSS * FSCONTROL* and *CROSS * CLCONTROL*, to Equation (10) when testing H1 and H2 does not alter our statistical inferences.

NBS and *NGS* as operationally or geographically diversified firms are more likely to have major plants, subsidiaries, or markets in foreign countries, and thus these firms may have greater incentives to cross-list in foreign countries. We include *BIG4* because cross-listed firms are less likely to hire non-Big 4 auditors due to their concerns over audit quality.

We use *SAG* (i.e., the sales growth) as an instrumental variable in the first-stage regression for the following reason. According to the analytical result in Observation 3, firms with greater funding requirements are more likely to be cross-listed in stronger legal regimes to allow their shareholders to enjoy the higher value of audits in spite of the compliance costs and increased audit fees. Therefore, we use the sales growth, which captures the firm's need for expansion and investment funding, as an important independent variable in the first-stage probit model. Furthermore, it has been well established in the audit fee literature that, at the firm level, the three major determinants of the audit fee are client size, client complexity, and client-specific risk. Hence, it is unlikely that the sales growth affects the audit fee via a channel other than the cross-listing decision. In other words, we believe that *SAG* is a valid exclusion restriction; therefore, we include *SAG* in the first-stage model but exclude it from the second-stage audit fee regression.

Given incomplete panel data of cross-listings, we estimate the probit cross-listing choice model using the average value of each variable for each firm over different years (except for *BIG4* and the time-invariant *REGIME*, *DISCL*, and country indicators). For *BIG4* we use the data of the first year that the firm appears in our data set (the use of the average value of *BIG4* does not alter the results). The explanatory power (pseudo R^2) of the model is 0.056 without country indicators and 0.155 with country indicators. In addition, we repeat the analysis using only the final year's data for every firm included in our data set. Regardless of which approach is used for the first-stage probit estimation, the results of the second-stage regressions with the firm-specific time-invariant inverse Mills ratios added remain qualitatively similar to the results reported in Tables 5 and 6.[23]

Alternative Measures of the Strength of Legal Regimes

As alternative measures of the strength of legal regimes, we consider the legal enforcement index (La Porta et al. 1997), the efficiency of the judicial system index (La Porta et al. 1998), the law and order index (La Porta et al. 1998), a composite score (which equals the investor rights score plus half of the law enforcement score) as used in Choi and Wong (2007) and Francis et al. (2003), and an instrumental variable approach (Leuz et al. 2003). The use of these alternative measures yields results qualitatively similar to the results that are based on the Wingate index, as these measures/indices are highly correlated with one another.

Other Sensitivity Analyses

Our analytical model assumes that the manager's reporting quality, captured by p, does not change with the cross-listing in stronger legal regimes. However, if the manager's reporting quality improves around the cross-listing, then the improved reporting quality has a negative effect on the expected legal liability cost and thus the audit fee (since $\frac{\partial f}{\partial p} < 0$). At the same time, as shown in Section II, an increase in the strength of the legal regime has a positive effect on the expected legal liability cost and thus the audit fee (since $\frac{\partial f}{\partial r} > 0$). Whether the audit fee increases with the cross-listing would depend on which of the above forces is dominant. In other words, the improved reporting quality would introduce a (conservative) bias against finding the positive effect of a stronger legal regime on the audit fee. Our empirical finding of a significant cross-listing fee premium suggests that the positive effect of the legal regime on the expected liability cost (and thus the audit fee) is the dominant factor. We maintain the assumption of a constant p to keep the model simple (without sacrificing the main insights of the model) and to focus on the dominant factor.

---

[23] We have also performed the following tests related to the endogeneity issue. First, we estimate the first-stage probit model separately for subsample H1 and subsample H2, and then perform the second-stage regressions accordingly. Second, we estimate the first-stage probit model with firm-specific variables and country indicators but without time-invariant variables (i.e., *DISCL* and *REGIME*). The results remain qualitatively similar.

Additional sensitivity tests are as follows. First, as the number of observations varies from 55 for Pakistan to 6,756 for the U.K. (see Table 2), the OLS results may be unduly influenced by the unequal size of the samples from different countries. To address this issue, we assign an equal weight to each country and estimate Equations (10) and (11) using the weighted least squares (WLS) procedure. The results are qualitatively identical to those in Tables 5 and 6. We also estimate Equations (10) and (11) without the U.K. observations, and find that the results remain qualitatively unchanged.

Second, we estimate Equations (10) and (11) by applying the Fama-MacBeth procedure. The results remain qualitatively unchanged, which suggests that the OLS results are robust to problems that may arise from residual cross-correlations.

Third, we estimate Equations (10) and (11) for each sample year and find that the year-by-year regression results are generally consistent with those reported in Tables 5 and 6. For example, the results of five out of the seven yearly regressions show that the coefficients on *CROSS * REGIME* in Equation (11) are significantly negative.

Finally, we perform a median quantile regression and a robust regression to minimize the influence of extreme observations without removing them from regression analyses. The coefficient estimates from these regressions are consistent with those in Tables 5 and 6.

VII. SUMMARY AND CONCLUDING REMARKS

This study examines the effects of cross-listings on audit quality and audit fees. Our analyses of the equilibrium audit fee predict that: (1) auditors charge higher fees for the firms that are cross-listed in stronger legal regimes than for non-cross-listed firms; (2) when firms are cross-listed in countries whose legal regimes are no stronger than those of their home countries, the cross-listing premium exists if, and only if, cross-listings lead to a significant increase in audit complexity; and (3) for the firms cross-listed in stronger legal regimes, the cross-listing audit fee premium increases (decreases) as the difference in the strength of legal regimes between the cross-listed foreign country and the home country becomes larger (smaller). We then test the model's predictions using a large sample of audit clients from 14 countries with different legal regimes.

The results of our cross-country, cross-firm regressions strongly support the first and third predictions, and are robust to a battery of sensitivity checks. In addition, we find that the cross-listing fee premiums are insignificant for the firms that are cross-listed in countries whose legal regimes are no stronger than those of their home countries. This indicates that the observed cross-listing premium is driven mainly by an upward shift in the strength of the legal regime caused by a cross-listing rather than by a change in audit complexity.

In addition, we show that a stronger legal regime induces higher audit quality, and thus shareholders are willing to pay the audit fee premiums to take advantage of the higher-value audits that arise from CL firms submitting themselves to stronger foreign legal regimes. Our study thus offers insights into the value of audits in a cross-listing context.

Our study has a number of limitations. First, we identify cross-listed firms based on foreign equity trading data from Worldscope, which does not distinguish between cross-listings on foreign stock exchanges and the other trading in foreign countries without cross-listings. Hence, our results may suffer from some noise. We argue, however, that this noise would introduce a conservative bias against rejecting the null hypothesis of no cross-listing effect, because the auditor's legal liability is likely to be greater for cross-listings through major stock exchanges than through other foreign trading such as private placements. Our sensitivity analysis using foreign firms traded in the U.S. supports this argument. Nevertheless, our cross-country results should be interpreted with caution.

Second, we use the firm-level and country-level variables together in one regression; hence, the t-statistics for the country-level variables and for the interaction variable, *CROSS * REGIME*, may be overstated. However, as *CROSS* is a firm-level variable, the coefficient on *CROSS* is not prone to such a bias.

Third, unlike previous single-country-based audit fee studies, our audit fee model is less complete in the sense that it is difficult to fully control for institutional differences across countries. Hence, our results could be biased to the extent that these omitted institutional factors are correlated with our test variables. Moreover, the explanatory power of our audit fee models is lower than is typical in single-country studies. The R^2 of our audit fee models is about 60 percent, while (for example) in DeFond et al. (2000), which is based on U.S. data, the R^2 is 66 percent.

Finally, although our sensitivity tests use the two-stage treatment effect approach to address endogenous cross-listing decisions, the explanatory power of the first-stage model is low (with pseudo R^2 of 0.155). Hence, our two-stage analysis may suffer from weak instrument problems. One cannot therefore rule out the possibility that the firms' endogenous cross-listing choices may influence our test results, and that this endogenous relation may differ across countries with different legal regimes. Nevertheless, the results of our cross-country regressions provide useful insights into the effects of cross-listings and the associated shift in legal regimes on the audit fee and the value of the audit, an issue that could not be effectively addressed in a single-country setting.

REFERENCES

Choi, J. H., and T. J. Wong. 2007. Auditor's governance function and legal environments: An international investigation. *Contemporary Accounting Research* 24 (1): 1–36.

———, J-B. Kim, X. Liu, and D. A. Simunic. 2008. Audit pricing, legal liability regimes, and Big 4 premiums: Theory and cross-country evidence. *Contemporary Accounting Research* 25 (2): 55–100.

Chung, R., M. Firth, and J. B. Kim. 2008. Big 6 conservatism around the world. Working paper, Hong Kong Polytechnic University.

Cook, R. D. 1977. Detection of influential observations in linear regression. *Technometrics* 19: 15–18.

DeFond, M. L., F. R. Francis, and T. J. Wong. 2000. Auditor industry specialization and market segmentation: Evidence from Hong Kong. *Auditing: A Journal of Practice & Theory* 19 (1): 49–66.

Doidge, C., G. A. Karolyi, and R. M. Stulz. 2004. Why are foreign firms listed in the U.S. worth more? *Journal of Financial Economics* 71 (2): 205–238.

Dye, R. A. 1995. Incorporation and the audit market. *Journal of Accounting and Economics* 19 (1): 75–114.

Fan, J., and T. J. Wong. 2005. Do external auditors perform a corporate governance role in emerging markets? Evidence from East Asia. *Journal of Accounting Research* 43 (1): 35–72.

Foerster, S. R., and G. A. Karolyi. 1999. The effects of market segmentation and investor recognition on asset prices: Evidence from foreign stock listing in the U.S. *The Journal of Finance* 54 (3): 981–1013.

Francis, J. 1984. The effect of audit firm size on audit prices. *Journal of Accounting and Economics* 6 (2): 133–151.

———, and D. Simon. 1987. A test of audit pricing in the small-client segment of the U.S. audit market. *The Accounting Review* 62 (1): 145–157.

———, D. Wang, and A. Nikitkov. 2003. The effect of legal environments on Big Five auditors' conservatism around the world. Working paper, University of Missouri.

Greene, W. 1997. *Econometric Analysis.* Upper Saddle River, NJ: Prentice Hall.

Hail, L., and C. Leuz. 2008. Cost of capital and cash flow effects of U.S. cross listings. *Journal of Financial Economics.*

Klapper, L., and I. Love. 2004. Corporate governance, investor protection and performance in emerging markets. *Journal of Corporate Finance* 10 (5): 703–728.

Lang, M., K. V. Lins, and D. Miller. 2003. ADRs, analysts, and accuracy: Does cross listing in the United States improve a firm's information environment and increase market value? *Journal of Accounting Research* 41 (2): 317–345.

———, J. Raedy, and M. Yetman. 2003. How representative are firms that are cross-listed in the United States? An analysis of accounting quality. *Journal of Accounting Research* 41 (2): 363–386.

La Porta, R., F. Lopez-de-Silanes, A. Schleifer, and R. W. Vishny. 1997. Legal determinants of external finance. *The Journal of Finance* 52 (3): 1131–1150.

———, ———, ———, and ———. 1998. Law and finance. *The Journal of Political Economy* 106 (6): 1113–1155.

———, ———, and ———. 2006. What works in securities laws? *The Journal of Finance* 61 (1): 1–32.

Lee, D. 2003. Why does shareholder wealth increase when foreign firms announce their listing in the U.S.? Working paper, University of Kentucky.

Leuz, C., D. Nanda, and P. Wysocki. 2003. Investor protection and earnings management: An international comparison. *Journal of Financial Economics* 69 (3): 505–527.

Miller, D. P. 1999. The market reaction to international cross-listings: Evidence from depository receipts. *Journal of Financial Economics* 51 (1): 103–123.

Pae, S., and S. Yoo. 2001. Strategic interaction in auditing: An analysis of auditor's legal liability, internal control system quality, and audit effort. *The Accounting Review* 76 (3): 333–356.

Pagano, M., A. Roell, and J. Zechner. 2002. The geography of equity listing: Why do companies list abroad? *The Journal of Finance* 57 (6): 2651–2694.

Reese, W. A., and M. S. Weisbach. 2002. Protection of minority shareholder interests, cross-listings in the United States and subsequent equity offerings. *Journal of Financial Economics* 66 (1): 65–104.

Sarkissian, S., and M. Schill. 2004. The overseas listing decision: New evidence of proximity preference. *Review of Financial Studies* 17 (3): 769–809.

Seetharaman, A., F. A. Gul, and S. G. Lynn. 2002. Litigation risk and audit fees: Evidence from U.K. firms cross-listed on U.S. markets. *Journal of Accounting and Economics* 33 (1): 91–115.

Shi, Y., M. Magnan, and J.-B. Kim. 2008. Why do cross-listed firms in the U.S. voluntarily release management earnings forecasts? Working paper, Concordia University.

Simunic, D. A. 1980. The pricing of audit services: Theory and evidence. *Journal of Accounting Research* 18 (1): 161–190.

Stulz, R. M. 1999. Globalization, corporate finance, and the cost of capital. *Journal of Applied Corporate Finance* 12 (3): 8–25.

Widder, D. V. 1989. *Advanced Calculus.* New York, NY: Dover Publications, Inc.

Wingate, M. 1997. An examination of cultural influence on audit environment. *Research in Accounting Regulation* 11 (Supplement): 129–148.

THE ACCOUNTING REVIEW
Vol. 84, No. 5
2009
pp. 1465–1493

American Accounting Association
DOI: 10.2308/accr.2009.84.5.1465

The Effects of Expected and Actual Accounting Choices on Judgments and Decisions

Shana M. Clor-Proell
University of Wisconsin–Madison

ABSTRACT: This research investigates how financial statement users' judgments and decisions are affected by the extent to which a firm's actual accounting choices match users' expectations. Based on prior communications research, I predict and find that users' credibility judgments are more extreme when a firm's actual accounting choices do not match expectations. Experiment 1 supports this prediction in a stock-based compensation context, and Experiment 2 supports it in an accounting estimate context. Further, evidence from Experiment 1 supports the prediction that credibility judgments mediate the effect of a mismatch on investment decisions. Finally, evidence from Experiment 2 partially supports the prediction that users who encounter a mismatch between actual and expected accounting choices are more likely to search for additional information than are users who encounter a match. The results have implications for accounting researchers, regulators, and managers interested in understanding how firms' accounting choices affect users' decisions.

Keywords: *accounting choices; accounting estimates; expectancy violations theory; stock-based compensation.*

Data Availability: *Contact the author.*

This research is based on my doctoral dissertation completed at Cornell University. I thank my dissertation committee members, Mark Nelson (Chair), Robert Bloomfield, Bob Libby, and Marty Wells, for their advice, support, and encouragement. This work has also benefited from helpful comments provided by two anonymous reviewers, Marty Butler, Bryan Church, Pat Hopkins, Kevin Jackson, Karla Johnstone, Steve Kachelmeier (editor), Kathryn Kadous, Susan Krische, Marlys Lipe, Laureen Maines, Derek Oler, Mark Peecher, Chad Proell, Kenny Reynolds, Kathy Rupar, Steve Salterio, Nick Seybert, Bill Tayler, Jane Thayer, Kristy Towry, Terry Warfield, and workshop participants at Cornell University, Emory University, Georgia Institute of Technology, Indiana University, University of Illinois, University of Utah, University of Wisconsin–Madison, and Queen's University. I thank the Deloitte & Touche Foundation and the S.C. Johnson Graduate School of Management at Cornell University for their generous financial support.

This research was awarded the AAA Accounting, Behavior & Organizations Section Outstanding Doctoral Dissertation Award in 2008.

Editor's note: Accepted by Steven Kachelmeier, with thanks to Marlys Lipe for serving as editor on a previous version.

Submitted: October 2007
Accepted: December 2008
Published Online: September 2009

I. INTRODUCTION

Firms often have discretion over accounting information presented in their financial statements. Managerial discretion can affect whether information is recognized or disclosed and can also affect the estimates underlying the accounting numbers. As users become accustomed to the discretion that firms have with respect to accounting information, they may develop expectations about how a firm treats various accounting transactions. These expectations can be based on industry norms or on a firm's past choices.

The purpose of this research is to examine how users' judgments and decisions are influenced by the extent to which a firm's accounting choices match users' expectations about those choices. Although prior accounting research examines users' preferences among accounting alternatives (e.g., Frederickson et al. 2006), it does not directly consider the role of users' expectations about accounting choices. Communications research on Expectancy Violations Theory (e.g., Burgoon and Burgoon 2001) suggests that expectations can affect how users interpret a firm's accounting choices. This theory further suggests that when there is a mismatch between actual and expected choices, users are more likely to seek additional information to find the reason for the mismatch.

I conduct two experiments to test my hypotheses. In Experiment 1, experienced accounting and finance practitioners receive information about a hypothetical firm, including information about its stock-based compensation under SFAS No. 123 (FASB 1995). The experiment uses a 2 × 2 × 2 between-subjects design, crossing actual accounting choice (recognition versus disclosure), whether actual accounting choice matches participants' expectations (match versus mismatch), and estimate uncertainty (high versus low). In addition to rating the reliability of the firm's reported net income, the firm's future profitability, and management's credibility, participants also make an investment decision.

Results from Experiment 1 show that the extent to which an actual accounting choice matches an expected accounting choice affects both credibility judgments and investment decisions, but does not affect assessments of future profitability and information reliability. Specifically, I find that when a firm chooses to recognize stock-based compensation expense, users view the firm's management to be more credible and invest more in the firm when there is a mismatch between actual and expected accounting choice than when there is a match. Conversely, when a firm chooses to disclose stock-based compensation expense, users view the firm's management to be less credible and invest less in the firm when there is a mismatch between actual and expected accounting choice than when there is a match. Further, credibility judgments mediate the relationship between the independent variables and the investment decision. Although I expected a mismatch would make participants more likely to acquire information about estimate uncertainty, I do not find a three-way interaction among actual accounting choice, match, and estimate uncertainty. This lack of interaction may be because all participants, not just those who viewed a mismatch, acquired the estimate-uncertainty information.

The purpose of Experiment 2 is to expand upon the results of Experiment 1 in three ways. First, it provides further evidence that the results of Experiment 1 are because of participants' expectations and not the result of an information signal associated with a firm's accounting choice. Specifically, it utilizes a control condition in which the information signal is the same as in the mismatch condition but expectations differ. Second, Experiment 2 extends to an accounting estimate setting the judgment results from Experiment 1 as a way to test the generality of the effects. Third, Experiment 2 uses alternative methods to investigate whether perceiving a mismatch is likely to result in an increased search for information.

In Experiment 2, students in a graduate-level financial statement analysis class rate a firm's net income reliability, future profitability, and management credibility. Participants also indicate how likely they are to seek additional information about the firm's fixed assets. They also choose specific questions they want answered about the firm. Experiment 2 employs a 1 × 3 between-subjects design that manipulates participants' expectations (match versus mismatch versus control). Results from Experiment 2 show that users react more strongly to a firm's choice when there is a mismatch. Furthermore, users react more strongly to a firm's choice when there is a mismatch than when there is no expectation, thus demonstrating that results are driven by expectations rather than by the information signal associated with the choice. As in Experiment 1, I find that expectations affect credibility assessments but do not affect assessments of information reliability or future profitability. Results also show that when there is a mismatch, users are more likely to seek additional information and may be more likely to request information specific to the firm's fixed assets.

This research has implications for accounting researchers, standard-setters, and regulators concerned with the judgments of nonprofessional investors. Specifically, I introduce the role of expectations as an important moderator in the context of accounting choice. This distinction is important because managers have discretion in a variety of accounting settings, and users may have distinct expectations about accounting alternatives. Furthermore, I find that expectations about accounting choices can affect users' propensity to search for additional information. Much prior accounting literature has been concerned with understanding the factors that affect information search and information acquisition of nonprofessional investors (e.g., Ackert et al. 1996, 1997). I contribute to this literature by demonstrating that nonprofessional investors' expectations may affect their information search behavior.[1] This finding may also be of interest to standard-setters and regulators concerned with the factors that affect the use of additional disclosures because it suggests that users are more likely to access additional disclosures if management has made an accounting choice that is inconsistent with users' expectations.

My research has practical implications for managers as evidence that users' expectations will affect how they respond to management's accounting choices. While much consideration has been given to how users' expectations affect interpretation of a firm's earnings (e.g., Bartov et al. 2002), little consideration has been given to how users' expectations affect interpretation of a firm's accounting choices. My findings suggest that managers could benefit from understanding the potential consequences of providing users with expectation-consistent versus inconsistent treatments. Instead of making choices consistent with its past choices or with the choices of other firms in its industry, a firm can distinguish itself either positively or negatively by making alternative choices.

The next section provides background information and describes my predictions. Sections III and IV describe the experiments used to test my predictions and present the results of each experiment, respectively. Section V discusses the implications.

II. BACKGROUND AND HYPOTHESES

Prior Research on Actual Accounting Choice

Research on accounting choice examines how market participants react to management's voluntary accounting choices. Prior research has examined a number of reporting contexts, including comprehensive income reporting (Hirst and Hopkins 1998; Maines and

---

[1] Prior research by Hirst et al. (1995) that considered the role of expectations in another setting did not consider how expectations could affect users' information search behavior.

McDaniel 2000), hybrid financial instruments (Hodge et al. 2006; Hopkins 1996), stock-based compensation (Frederickson et al. 2006; Hodge et al. 2004), segment reporting (Maines et al. 1997), and consolidations (Hopkins et al. 2000).

Research finds that managers' voluntary accounting choices influence how users acquire, interpret, and weight information. These studies have found that both sophisticated and unsophisticated users may have difficulty acquiring financial information because of differences in accounting treatment. For example, Hirst and Hopkins (1998) find that analysts are more likely to attend to information placed in the income statement than to information placed in the statement of stockholders' equity. Similarly, Hodge et al. (2004) find that nonprofessional investors are less likely to acquire stock-based compensation information if it is placed in a footnote rather than in the income statement and, therefore, invest more in a firm that discloses than in a firm that recognizes stock-based compensation.

Even in circumstances in which users are able to acquire accounting information despite differences in treatment, accounting choices can still affect users' evaluation and weighting of information via the signal that a choice sends. For example, Maines and McDaniel (2000) argue that information placement can convey a signal about the relevance of information, with information that is disclosed being perceived as less useful than information that is recognized. Similarly, Frederickson et al. (2006) argue that discretionary accounting choices can convey information about a firm's future profitability, with income-decreasing accounting choices signaling that managers expect higher future profitability. This assertion is consistent with prior archival findings that the market responds favorably to firms that elect voluntary expense recognition (Aboody et al. 2004).

Expectations about Accounting Choice and Users' Judgments

Although prior research provides evidence on users' reactions to managers' voluntary accounting choices, it generally is silent on the role of users' expectations about these choices. Communications research on Expectancy Violations Theory (EVT) assumes that people form expectations regarding the communications they receive. EVT is primarily concerned with what happens when expectations about communications are violated, meaning that the communication does not occur in the way it was anticipated by the recipient. The theory posits that the violation will catch the recipient's attention and that the recipient consequently will respond more strongly to the violation. In other words, positive violations will result in more favorable reactions than communications that conform to expectations, but negative violations will result in less favorable reactions than communications that conform to expectations. Furthermore, the theory posits that the recipient will try to make sense of the violation and its implications by engaging in deeper information processing (Burgoon and Burgoon 2001). Prior EVT research has found that expectancy violations can affect assessments of the communicator's credibility, attractiveness, and influence (Burgoon and Burgoon 2001).[2]

Prior accounting research provides evidence that accounting choices affect perceptions of management credibility. Hodge et al. (2006) examine perceptions of management credibility when management chooses to classify hybrid securities as either debt or equity. In that paper, management's discretionary choice alters the appearance of the firm's performance; classifying the securities as debt makes the firm appear closer to violating a debt covenant. Hodge et al. (2006) find that users perceive management as more credible if management chooses the incentive-inconsistent treatment (debt) than if it chooses the incentive-consistent treatment (equity). Taken together, the accounting literature demonstrates that accounting choices affect credibility assessments and EVT suggests that expectations will also affect these assessments. Thus, I form the following hypothesis about how a mismatch between actual and expected accounting choices will affect credibility judgments, represented graphically in Figure 1, Panel A, for Experiment 1:

---

[2] EVT has been tested in numerous settings, and empirical tests generally support the theory (e.g., Bevan 2003; Burgoon 1983; Burgoon and Hale 1988; Burgoon and LePoire 1993; Burgoon et al. 1989; Campo et al. 2004; Campo and Cameron 2006; Houser 2005; LePoire and Burgoon 1994).

> **H1:** The effect of an accounting choice on users' credibility assessments is more extreme when there is a mismatch between actual and expected accounting choice than when there is a match between actual and expected accounting choice.

Prior accounting research suggests that a mismatch could affect additional judgments specific to the accounting setting. Specifically, the interaction described above might also affect the perceived reliability of a choice or the implications of a choice for a firm's future profitability. Frederickson et al. (2006) find that in discretionary settings, users perceive net income to be more reliable when management chooses to recognize rather than merely disclose costs. Similarly, users view the firm as most attractive when management voluntarily chooses recognition and least attractive when it voluntarily chooses disclosure. Both Frederickson et al. (2006) and Aboody et al. (2004) assert that this difference arises from the signal that voluntary expense recognition conveys about future profitability, but they do not directly test this hypothesis. If expectations about accounting choices affect credibility judgments, then it is possible that expectations will also affect reliability and future profitability judgments. Although I do not make explicit hypotheses about these judgments, I pose the following research question:

> **RQ1:** Are assessments of reliability and future profitability affected by a mismatch between actual and expected accounting choices?

Taken together, the findings of Hodge et al. (2006), Frederickson et al. (2006), and Aboody et al. (2004) suggest that the effect of actual and expected accounting choices are likely to affect assessments of management credibility, reliability, and future profitability. These studies may also be picking up the same underlying construct. I simultaneously elicit multiple judgments to determine the extent to which expectations affect multiple constructs.

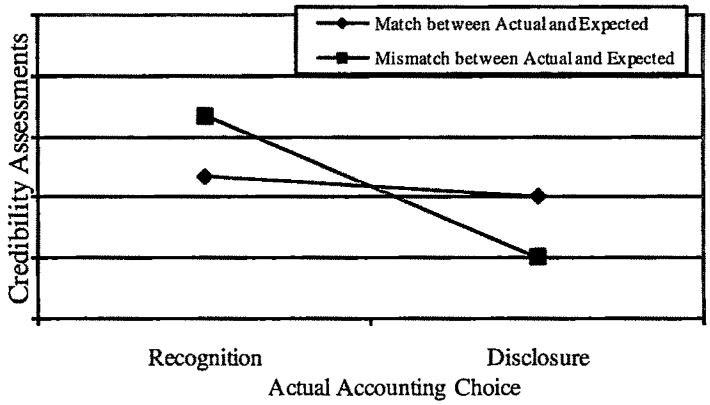
Expectations about Accounting Choice and Users' Investment Decisions

Management research by Chatterjee et al. (1999) and Allessandri and Khan (2006) has argued that firms' actions, including their accounting choices, can become institutionalized, thus creating expectations, and that a firm's market performance is affected when managers choose strategies that do not match expectations. This assertion is consistent with prior research examining market response to discretionary accounting choices, and can be interpreted in the context of EVT. Specifically, Aboody et al. (2004) examine firms that voluntarily elect to recognize stock-based compensation expense. They find positive and significant abnormal returns for early announcers but insignificant abnormal returns for late announcers.

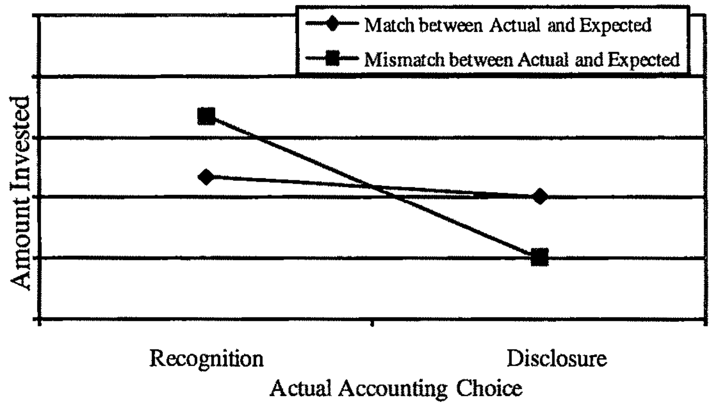
This result could be because recognition for early announcers was unexpected because the norm had been for most firms to disclose. Therefore, recognition could have constituted

FIGURE 1
Experiment 1 Predictions

Panel A: Interaction between Actual Accounting Choice and Match on Credibility Assessments[a]



Panel B: Interaction between Actual Accounting Choice and Match on Investment Decisions[b]



---

[a] Panel A displays the two-way interaction described in H1. Participants rate management's credibility on a nine-point scale. Larger numbers indicate higher credibility assessments.

[b] Panel B displays the two-way interaction described in H2. Participants indicate how much they would like to invest in the firm. Participants in the actual recognition (disclosure) condition receive information indicating that the target firm has elected to recognize (disclose) its stock-based compensation. When there is a match (mismatch) between actual and expected accounting choice, participants are informed that the target firm has elected the same (the opposite) treatment as other firms in the industry.

a positive expectancy violation because recognition is viewed more favorably than disclosure (Frederickson et al. 2006). However, recognition was expected for the late announcers and did not constitute an expectancy violation, because many firms had already adopted recognition. This result implies that users respond differently to a firm's accounting choices, depending on their expectations.[3] Based on this research, I predict that the interaction between actual accounting choice and the extent to which it matches expectations will affect investment decisions, as is represented graphically in Figure 1, Panel B, for Experiment 1:

> **H2:** The effect of an accounting choice on users' investment decisions is more extreme when there is a mismatch between actual and expected accounting choice than when there is a match between actual and expected accounting chóice.

Furthermore, because users' investment decisions are based on their judgments about the firm, it logically follows that the judgments users make based on actual and expected accounting choices mediate users' investment decisions. Thus, I make the following hypothesis and examine the following research question:

> **H3:** Assessments of management credibility mediate the effects of actual accounting choice and match/mismatch on users' investment decisions.
>
> **RQ2:** Do assessments of reliability and future profitability mediate the effects of actual accounting choice and match/mismatch on users' investment decisions?

Figure 2 shows the relationships among the primary independent variables, judgments, and decisions examined in Experiment 1.

Expectations about Accounting Choice and Search Behavior

As discussed above, EVT predicts a stronger reaction to an unexpected choice. EVT then goes a step further by arguing that when expectations are violated, the receiver will try to make sense of the violation by seeking additional information. In other words, when there is a mismatch between actual and expected accounting choice, EVT predicts that users engage in more extensive processing by searching for and/or weighting information that explains the mismatch. When actual and expected accounting choices match, EVT predicts that users will engage in less extensive cognitive processing. Therefore, they are less likely to search for and incorporate additional information into their judgment. I test this reasoning in the following hypothesis:

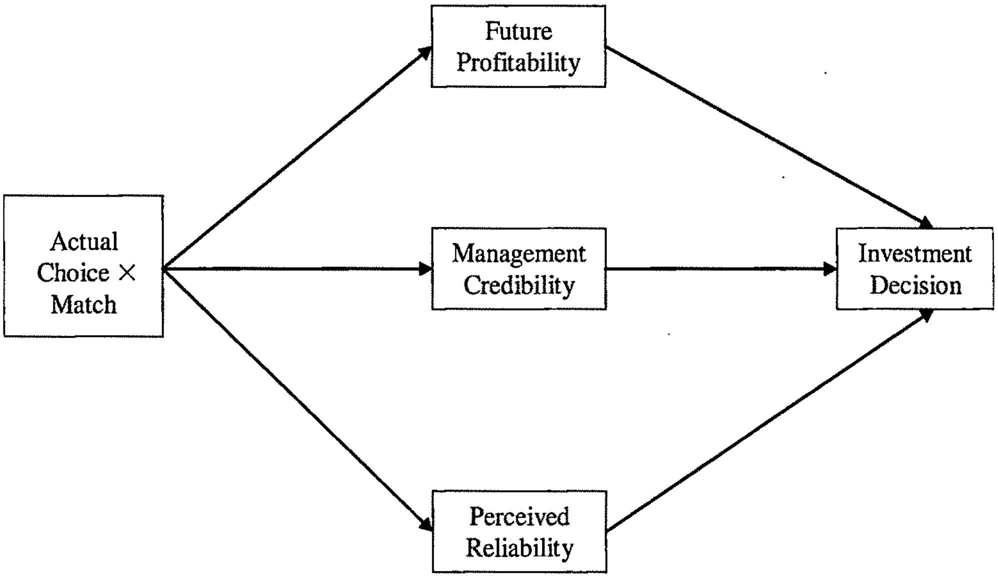
> **H4:** A mismatch between actual and expected accounting choices results in greater information search than does a match between actual and expected accounting choices.

In Experiment 1, I manipulate estimate uncertainty information for participants to search for as an explanation for a mismatch. I expect that participants who receive a mismatch are more likely to search for and incorporate differences in estimate uncertainty information into their judgment than are participants who receive a match. In Experiment

---

[3] Aboody et al. (2004) could not distinguish whether their results were because of the effect of expectations based on industry norms or the result of two alternative explanations: (1) the initial hype surrounding voluntary recognition was later found to be unjustified, and (2) early adoption provides a signal about the firm's future prospects that late adoption does not provide. My research untangles these explanations by examining the effect of expectations while controlling all other information available to users and eliciting perceptions of the firm's prospects. Although I do not provide evidence that expectations held for early announcers differ from expectations held for late announcers, I do provide evidence that expectations affect the reaction to accounting choice.

FIGURE 2
Experiment 1
Relationships among the Primary Independent Variables, Judgments, and Decisions



The interaction between actual accounting choice and match may directly affect perceptions of future profitability, management credibility, and/or net income reliability. In addition, the interaction may affect investment decisions via perceptions of future profitability, management credibility, and/or net income reliability.

2, I ask participants how likely they are to search for more information, and I examine the type of information they want to obtain.

III. EXPERIMENT 1

Stock-Based Compensation

Experiment 1 tests the above predictions and research questions in the context of stock-based compensation. Before 1995, APB Opinion No. 25 (APB 1972) required firms to use the intrinsic-value method for stock options. This method recognized expense as equal to the excess on the measurement date (generally, the grant date) of the market value of the options over the option exercise price. Because this amount was typically zero, no expense was recorded in the income statement.

A 1993 exposure draft of SFAS No. 123 (FASB 1995) required firms to use the fair-value method to recognize stock-based compensation expense in the income statement. This method uses an option-pricing model to estimate the value of the options at the grant date. This value is then amortized as expense over the vesting period of the options. Due to intense political pressure, the final version of SFAS No. 123 (FASB 1995) allowed firms to calculate recognized stock-based compensation expense using either the intrinsic-value or the fair-value method. Firms that elected the intrinsic-value method were required to disclose in the footnotes what the stock-based compensation expense would have been if the fair-value method been used. Although the FASB encouraged the use of the fair-value method, nearly all firms initially chose the intrinsic-value method. After numerous scandals in the late 1990s and in anticipation of the issuance of SFAS No. 123(R) (FASB 2004), many firms switched to the fair-value method (Aboody et al. 2004). SFAS No. 123(R), effective for accounting periods after June 15, 2005, mandates use of the fair-value method.

I use the stock-based compensation context in Experiment 1 for several reasons. First, given recent changes in standards, users are likely to understand the basic accounting for stock-based compensation. Furthermore, because option-pricing models differ in the uncertainty associated with the resulting expense, this setting should allow me to examine whether a mismatch affects acquisition of estimate uncertainty information. Finally, several studies of accounting choice have employed this context, so using this context enhances the comparability of my results to those of past research.

Method

Participants

Five hundred sixty-six alumni from a top-20 M.B.A. program who work in accounting- and finance-related fields were selected from an alumni database and asked via email to participate in the experiment. A total of 172 alumni (30 percent) participated during October 2006 and were randomly assigned to the experimental treatments. Participants averaged 10 years of work experience, had taken an average of 4.63 accounting classes, and were an average of 35 years old. Eighty-three percent are male.[4]

Design

The experiment utilizes a 2 × 2 × 2 design, varying actual accounting choice (recognition versus disclosure), whether actual choice matches users' expectations (match versus mismatch), and estimate uncertainty (high versus low). Actual accounting choice is manipulated by varying whether the target firm has chosen to include the fair value of stock-based compensation as an expense on the income statement or only in the footnotes. Match is manipulated by informing participants that the target firm's choice is either consistent with or inconsistent with the choice of most firms in the industry. Estimate uncertainty is manipulated by varying the extent to which the target firm can accurately estimate the assumptions in the Black-Scholes model. The Appendix contains the wording used to manipulate estimate uncertainty.

Materials and Procedure

The experiment is administered via a six-part web-based instrument. Participants are required to answer all questions in each section and, except as noted below, are not allowed to return to a previous section after its completion. Figure 3 summarizes the information provided in each section of the materials.

Section 1 informs participants that their task is to act as a potential investor trying to decide how to allocate $10,000 between two firms in the same industry. The background information, based on a case used by Nelson et al. (2005), informs participants that Capital Auto Parts, Inc. (CAP) and Parts, Inc. are both medium-sized automobile parts manufacturers. Although CAP is the target firm of interest in the experiment, Parts, Inc. is provided to create a relative investment decision, which limits noise in my dependent variable. Participants are also provided with a brief example of accounting for stock-based compensation

---

[4] I also collected information about the native language spoken by participants and about participants' current job function. I found that 78 percent of my participants considered English their first language. Thirteen percent worked in asset management, 21 percent in financial analysis, 14 percent in investment banking, 17 percent in managerial finance, and 35 percent in a variety of other areas. Except as discussed elsewhere, none of the demographic variables were significant in any of my analyses, so they will not be discussed further.

FIGURE 3
Experiment 1 Summary

Section 1	Section 2	Section 3	Section 4	Section 5	Section 6
Task Description	Actual Accounting Choice Manipulation	Industry Information	Investment Question *or* Mediator Measures	Investment Question *or* Mediator Measures	Debriefing Questions
Relevant Accounting Standard		EPS for CAP and Parts, Inc.			
Expected Accounting Choice Manipulation		Footnotes (includes Estimate Uncertainty Manipulation)			

The above table summarizes the information provided to participants in each section of the experimental materials. The order in which the mediators are elicited relative to the investment decision is balanced so that half the participants respond to the mediators in Section 4, and the other half of the participants respond to the investment decision in Section 4.

under APB Opinion No. 25 (APB 1972), SFAS No. 123 (FASB 1995), and SFAS No. 123(R) (FASB 2004). They are asked to assume that in all prior years, APB Opinion No. 25 was the prevailing standard. Beginning with the most recent fiscal year, the instructions ask participants to assume that the prevailing standard has changed to SFAS No. 123, such that management now has discretion as to whether to recognize stock-based compensation in net income or to disclose stock-based compensation in pro forma net income. Therefore, the setting is one in which the current year is the first in which firms have had this choice. In the final part of Section 1, participants are provided with additional information that indicates how most firms in the industry treated stock-based compensation in the most recent fiscal year. Participants are told that they expect CAP and Parts, Inc. to choose the treatment used by most firms in the industry. Explicitly setting participants' expectations is consistent with the approach used in past research (e.g., Burgoon and LePoire 1993).[5]

After responding to a manipulation check in which participants indicate which accounting choice they expect, participants move to Section 2, which reveals the method that CAP and Parts, Inc. actually chose. Parts, Inc.'s treatment is always consistent with the treatment chosen by the industry. CAP's treatment varies, based on experimental conditions, so that half the time it matches the industry and half the time it does not. After reading about CAP and Parts, Inc.'s actual choices, participants answer manipulation-check questions about the actual choice that each firm made and about whether CAP's choice is or is not consistent with that of Parts, Inc. and the industry. Participants also indicate their degree of surprise at CAP's treatment, based on the information provided.

---

[5] Expectations are set before participants know the actual accounting choice and estimate uncertainty so as to be consistent with most real-world settings and to minimize the extent to which estimate uncertainty information could affect participants' expectations about accounting choice.

Section 3 contains selected financial information for CAP and Parts, Inc. CAP and Parts, Inc. have similar assets per share, both have a similar trailing P/E ratio, and are of similar size. In addition, both CAP and Parts, Inc. have a similar amount of stock-based compensation. Next, participants receive a table that shows CAP and Parts, Inc.'s reported and pro forma net income for the current and prior fiscal year, which varies by treatment. The table is constructed to facilitate comparison of recognized and disclosed numbers to avoid information accessibility problems. Specific EPS numbers for participants are shown in the Appendix.

The final information in Section 3 includes four footnotes from CAP's most recent annual report. The first two footnotes are held constant across conditions and serve to give participants extraneous information as they try to determine the reason for CAP's reporting choice. By making the estimate-uncertainty information slightly harder to find, it is more likely to be accessed by participants in the mismatch condition than by participants in the match condition. The last two footnotes relate to stock-based compensation. The first footnote indicates how CAP has chosen to treat stock-based compensation, while the second provides information about the uncertainty of the estimates underlying CAP's stock-based compensation.

After reviewing this information, participants move to Sections 4 and 5, in which they provide their judgments and investment decision. While answering these questions, participants are able to click on a link that provides them with financial information from Section 3. The order in which the mediators are elicited relative to the investment decision is balanced such that half the participants respond to the mediators in Section 4, and the other half respond to the investment decision in Section 4.[6,7]

To provide further assurance that the measurement process does not automatically induce a relationship between mediators and investment decisions, I include an artificial mediator question about the legal liability associated with CAP's reporting decision. Because the accounting standard provided in the materials allows either recognition or disclosure, I do not expect the firm's choice to raise legal liability concerns. Finding that this mediator is not significantly correlated with the investment decision assures me that any significant relationships between the other mediators and the investment decision are genuine.

Finally, participants are directed to the debriefing questionnaire in Section 6, which contains demographic questions. Submitting this section concludes the experiment.

Dependent Variables

Participants make all judgments using nine-point scales. To measure perceived future profitability, participants indicate what CAP's reporting choice implies about its future performance. The endpoints of the scale are labeled "Very Unfavorable" and "Very Favorable," and the midpoint is labeled "Neutral." To measure perceived reliability, participants are informed that the FASB defines reliable information as information that is verifiable, reflects a business' activities in an unbiased manner, and is measured with little uncertainty (consistent with Frederickson et al. 2006). They assess the reliability of CAP's reported net

---

[6] The order in which future profitability, reliability, and credibility are elicited within their section is randomized.

[7] Manipulating the order in which participants respond to the mediators and the investment decision does not preclude prior responses from affecting later responses. However, it does allow me to examine the effect of the independent variables on the investment decision without the intrusion of the mediators for half the data. Similarly, it allows me to examine the effect of the independent variables on the mediators without the intrusion of the investment decision for half the data. The order in which the mediators and investment decision are elicited is not significant in any analysis and is not discussed further.

income with the endpoints of the scale labeled "Very Unreliable" and "Very Reliable" and the midpoint labeled "Somewhat Reliable."

To measure perceived management credibility, participants are informed that reporting credibility is defined as management's trustworthiness and competence in financial disclosure (consistent with Griffin [1967], Hovland et al. [1953], and Mercer [2005]). They rate CAP's reporting credibility relative to that of the industry with the endpoints of the scale labeled "Much Lower" and "Much Higher" and the midpoint labeled "Same."[8] To measure legal liability concerns, participants indicate the extent to which these concerns drove CAP's treatment of stock-based compensation, with the endpoints of the scale labeled "Not at All" and "Completely" and the midpoint labeled "Somewhat."

For the investment decision, participants decide how much of $10,000 they would like to invest in CAP. They are informed that any money not invested in CAP is automatically invested in Parts, Inc. They can choose an amount between zero and $10,000 in $1,000 increments.

Results

Manipulation Checks

Responses to the manipulation-check questions indicate that all manipulations are successful. Ninety-eight percent (100 percent) of the participants in the Actual Recognition (Disclosure) condition correctly indicate that CAP chose to recognize (disclose) its stock-based compensation in 2005. Excluding the participants who failed this manipulation-check question does not affect the results. Therefore, analyses include these participants.

To verify that the match manipulation succeeded, I examine whether participants viewed a mismatch between actual and expected accounting choice as less expected and more surprising than a match. Participants rated the unexpectedness of CAP's choice on a nine-point scale, with 1 (9) labeled "Completely Expected" ("Completely Unexpected") and the midpoint labeled "Somewhat Expected." Participants rated CAP's choice as significantly less expected ($t = 11.00$, $p = 0.00$) if they received a mismatch (mean = 5.22) than if they received a match (mean = 1.95).[9] Similarly, participants rated the surprise associated with CAP's choice on a nine-point scale, with 1 (9) labeled "Not at All Surprising" ("Very Surprising") and the midpoint labeled "Somewhat Surprising." Participants rated CAP's choice as significantly more surprising ($t = 14.97$, $p = 0.00$) if they received a mismatch (mean = 6.23) than if they received a match (mean = 2.36).

Participants rated CAP's confidence in the accuracy of its stock-based compensation estimate on a nine-point scale, with 1 (9) labeled "Not at All Confident" ("Very Confident") and the midpoint labeled "Somewhat Confident." Participants rated CAP as significantly more confident ($t = 20.23$, $p = 0.00$) in the low-estimate-uncertainty condition (mean = 7.57) than in the high-estimate-uncertainty condition (mean = 2.55). This indicates that the manipulation of estimate uncertainty was also successful.

Hypothesis Tests

Hypothesis 1 predicts that users' credibility judgments are more extreme when a mismatch occurs between actual and expected accounting choices than when a match occurs. Average credibility ratings are shown in Table 1, Panel A, with match cells unshaded and mismatch cells shaded. As shown in Table 1, Panel B, and in Figure 4, Panel A, an ANOVA

---

[8] Prior research generally elicits measures of credibility by using multiple questions. In Experiment 1, I elicited credibility by using only one question. This potential limitation is remedied in Experiment 2.

[9] All t-tests and associated p-values reported in the paper are associated with directional hypotheses and, therefore, are one-tailed.

TABLE 1
Experiment 1 Result
Average Credibility Ratings

Panel A: Cell Means[a]

	High Estimate Uncertainty				Low Estimate Uncertainty				Collapsed across Uncertainty	
Actual	**Expected Accounting Choice**				**Expected Accounting Choice**				**Expected Accounting Choice**	
Choice	**n**	**Recog.**	**n**	**Discl.**	**n**	**Recog.**	**n**	**Discl.**	**Recog.**	**Discl.**
Recog.	21	4.57	21	6.29	21	5.48	22	6.77	5.02	6.53
Discl.	21	3.57	22	4.36	22	4.18	22	5.14	3.88	4.75

Panel B: Test of H1[b]

Source	d.f.	S.S.	M.S.	F-value	p-value
Actual	1	92.01	92.01	47.66	0.00
Match	1	4.29	4.29	2.22	0.14
Actual * Match	1	60.80	60.80	31.49	0.00
Uncertainty	1	20.68	20.68	10.71	0.00
Actual * Uncertainty	1	0.00	0.00	0.00	0.99
Match * Uncertainty	1	0.90	0.90	0.47	0.50
Actual * Match * Uncertainty	1	0.18	0.18	0.09	0.76

---

[a] Panel A provides the mean credibility assessment associated with each treatment. Participants are informed that reporting credibility is defined as managements' trustworthiness and competence in financial disclosure. They are asked to rate CAP's reporting credibility relative to that of the industry on a nine-point scale with endpoints labeled "Much lower" and "Much higher" and a midpoint labeled "Same." Therefore, larger numbers indicate higher credibility assessments. Participants in the actual recognition (disclosure) condition receive information indicating that the target firm has elected to recognize (disclose) stock-based compensation. Participants in the expected recognition (disclosure) condition receive information indicating that most firms in the industry have chosen to recognize (disclose) stock-based compensation. When there is a match (mismatch) between actual and expected choice, participants are informed that the target firm has elected the same (opposite) treatment as other firms in the industry. In the low (high) estimate uncertainty condition, participants are informed that the target's management is (is not) very confident in the accuracy of its estimate. Shaded cells indicate conditions in which there is a mismatch between actual and expected accounting choice.

[b] Panel B provides the hypothesis test in which the mean credibility assessment is the dependent variable.

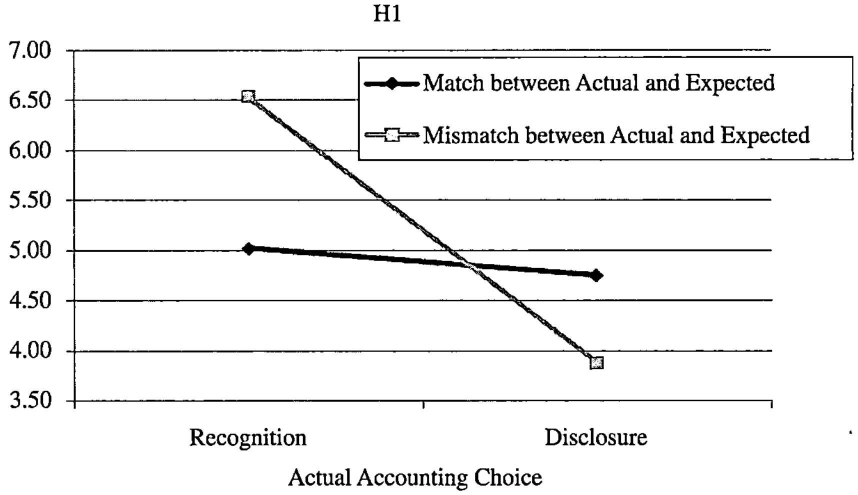
with credibility rating as the dependent variable supports H1. A significant interaction between actual accounting choice and match (F = 31.49, p = 0.00) indicates that the effect of an actual choice is greater when there is a mismatch between actual and expected accounting choice than when there is a match. Simple-effects tests indicate that credibility ratings in the recognition condition are significantly greater when participants receive a mismatch than when they receive a match (t = 4.99, p = 0.00), and credibility ratings in the disclosure condition are significantly lower when participants receive a mismatch than when they receive a match (t = 2.93, p = 0.00).

Hypothesis 2 predicts that users' investment decisions are more extreme when a mismatch occurs between actual and expected accounting choices than when a match occurs. Mean investment amounts are shown in Table 2, Panel A, with match cells unshaded and mismatch cells shaded. As shown in Table 2, Panel B, and in Figure 4, Panel B, an ANOVA with investment amount as the dependent variable supports H2. A significant interaction between actual accounting choice and match (F = 5.24, p = 0.02) indicates that the effect

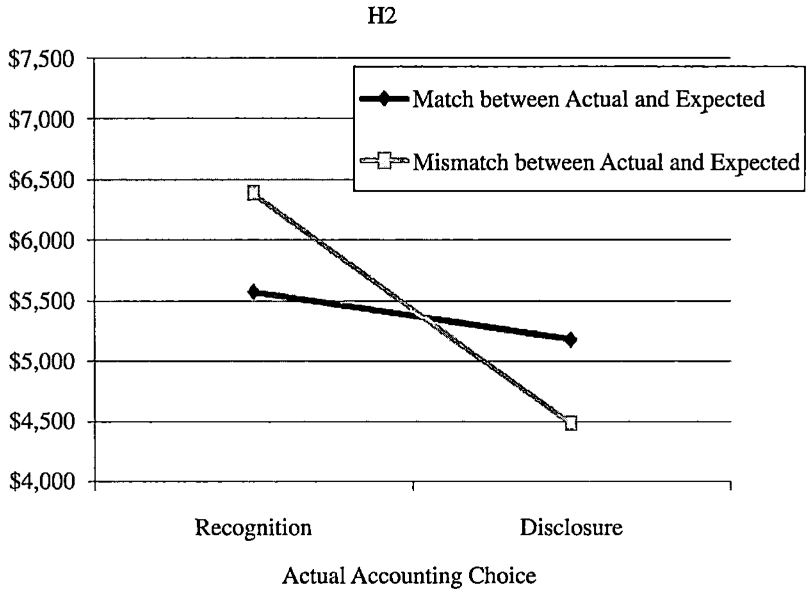
---

FIGURE 4
Experiment 1 Results

Panel A: Average Credibility Ratings[a]



Panel B: Amount Invested in CAP[b]



---

[a] Panel A displays the average credibility ratings based on a nine-point scale. Larger numbers indicate higher credibility assessments.

[b] Panel B displays the average investment amounts. Participants indicate how much they would like to invest in the firm in $1,000 increments ranging from $0 to $10,000. Participants in the actual recognition (disclosure) condition receive information indicating that the target firm has elected to recognize (disclose) stock-based compensation. When there is a match (mismatch) between actual and expected choice, participants are informed that the target firm has elected the same (opposite) treatment as other firms in the industry.

TABLE 2
Experiment 1 Results
Amount Invested in CAP

Panel A: Cell Means, (Medians), and [Standard Deviations] in Dollars[a]

	High Estimate Uncertainty				Low Estimate Uncertainty				Collapsed across Uncertainty	
	Expected Accounting Choice				Expected Accounting Choice				Expected Accounting Choice	
Actual Choice	n	Recog.	n	Discl.	n	Recog.	n	Discl.	Recog.	Discl.
Recog.	21	4,952	21	5,905	21	6,190	22	6,864	5,571	6,384
		(5,000)		(6,000)		(6,000)		(7,000)	(5,000)	(6,000)
		[2,459]		[1,895]		[1,834]		[2,122]	[2,258]	[2,049]
Discl.	21	4,048	22	4,864	22	4,909	22	5,500	4,478	5,182
		(3,000)		(5,000)		(5,000)		(5,000)	(5,000)	(5,000)
		[3,074]		[1,833]		[1,659]		[2,198]	[2,463]	[2,002]

Panel B: Test of H2 and H4[b]

Source	d.f.	S.S.	M.S.	F-value	p-value
Actual	1	56,612,315	56,612,315	12.00	0.00
Match	1	128,373	128,373	0.03	0.87
Actual * Match	1	24,700,506	24,700,506	5.24	0.02
Uncertainty	1	36,668,793	36,668,793	7.78	0.01
Actual * Uncertainty	1	1,312,909	1,312,909	0.28	0.60
Match * Uncertainty	1	7,865	7,865	0.00	0.97
Actual * Match * Uncertainty	1	683,190	683,190	0.14	0.70

[a] Panel A provides the mean investment amount, median investment amount, and standard deviation associated with each treatment. Participants indicate the amount they would like to invest in the firm in $1,000 increments, ranging from $0 to $10,000. Participants in the actual recognition (disclosure) condition receive information indicating that the target firm has elected to recognize (disclose) stock-based compensation. Participants in the expected recognition (disclosure) condition receive information indicating that most firms in the industry have chosen to recognize (disclose) stock-based compensation. When there is a match (mismatch) between actual and expected choice, participants are informed that the target firm has elected the same (opposite) treatment as other firms in the industry. In the low (high) estimate uncertainty condition, participants are informed that the target's management is (is not) very confident in the accuracy of its estimate. Shaded cells indicate conditions in which there is a mismatch between actual and expected accounting choice.

[b] Panel B provides the hypothesis tests in which the mean investment amount is the dependent variable.

of an actual choice is greater when there is a mismatch between actual and expected accounting choice than when there is a match. Simple-effects tests indicate that investment amounts in the recognition condition are significantly greater when participants receive a mismatch than when they receive a match ($t = 1.73$, $p = 0.04$), and investment amounts in the disclosure condition are significantly lower when participants receive a mismatch than when they receive a match ($t = 1.47$, $p = 0.07$).[10]

---

[10] Of all the demographic variables collected, none of these variables significantly predicts the investment decision. Only job function is marginally significant ($F = 2.97$, $p = 0.09$) when included in a model with the independent variables. The inclusion of job function does not affect the results presented in Table 2, Panel B.

To test H3, I investigate whether perceptions of management credibility mediate the relationship between the independent and dependent variables.[11,12] I conduct all mediation analyses according to the four-step procedure specified by Baron and Kenny (1986) and Kenny et al. (1998). Table 3 summarizes results from the four-step procedure used to investigate mediation of H2, the actual accounting choice by match interaction.

Regarding Step 1, the interaction between actual accounting choice and match is significantly correlated with the investment decision ($F = 5.24$, $p = 0.02$). Regarding Step 2, the interaction between actual accounting choice and match is significantly correlated with management credibility ($F = 31.49$, $p = 0.00$) but is not significantly correlated with either perceived reliability or future profitability. This result indicates that only management credibility is a potential mediator of the effect of the actual accounting choice by match interaction on the investment decision.

Regarding Step 3, I find a significant relationship between management credibility and the investment decision ($F = 34.59$, $p = 0.00$). Finally, when controlling for management credibility, the interaction between actual accounting choice and match becomes nonsignificant ($F = 0.04$, $p > 0.50$), and the effect of management credibility remains significant ($F = 40.12$, $p = 0.00$). This result satisfies Step 4 in the Kenny et al. (1998) procedure

TABLE 3
Experiment 1: Test of H3 and Analysis of Research Questions

Mediation of Actual Accounting Choice × Match Interaction

	F-value (p-value)			
	Step 1: IV Effect on DV	**Step 2: IV Effect on Mediator**	**Step 3: Mediator Effect on DV**	**Step 4: IV Effect on DV with Mediator**
IV: Actual Choice × Match				IV: 0.04
DV: Amount Invested	5.24	31.49	34.59	(0.83)
Mediator: Credibility	(0.02)	(0.00)	(0.00)	Mediator: 40.12
				(0.00)
IV: Actual Choice × Match				IV: 4.06
DV: Amount Invested	5.24	1.13	8.07	(0.05)
Mediator: Reliability	(0.02)	(0.29)	(0.01)	Mediator: 5.43
				(0.02)
IV: Actual Choice × Match				IV: 4.14
DV: Amount Invested	5.24	2.51	0.26	(0.04)
Mediator: Future Profitability	(0.02)	(0.12)	(0.61)	Mediator: 21.58
				(0.00)

This table summarizes the mediation tests for the actual accounting choice by match interaction. Management credibility fully mediated the interaction, but perceptions of future profitability and information reliability did not.

---

[11] Consistent with my expectations, the relationship between concerns about legal liability and the investment decision is not significant ($F = 2.09$, $p = 0.15$).

[12] Before conducting the mediation analyses, I verified by testing for collinearity that perceptions of future profitability, reliability, and management credibility were separate constructs. Common indicators of collinearity include the variance inflation factor (VIF), tolerance, and condition index. In general, collinearity is a potential problem if VIF values are greater than 10, tolerance values are less than 0.1, or condition indices are greater than 30. I find that all VIF values are less than 1.5, all tolerance values are greater than 0.7, and all condition indices are less than 10.2. Therefore, the three items do not appear to be collinear (Kleinbaum et al. 1998).

and indicates that, consistent with H3, management credibility fully mediates the effect of the actual accounting choice by match interaction on the investment decision.

Hypothesis 4 predicts that participants who receive a mismatch are more likely to seek additional information. Thus, I expect that differences in estimate uncertainty will have a greater effect on credibility judgments and investment decisions when there is a mismatch. Therefore, in Experiment 1, H4 is tested as a three-way interaction involving actual accounting choice, match, and estimate uncertainty. As shown in Table 1, Panel B, the three-way interaction does not significantly predict credibility ratings ($F = 0.09$, $p > 0.50$). In addition, as shown in Table 2, Panel B, the three-way interaction does not significantly predict investment decisions ($F = 0.14$, $p > 0.50$). Therefore, H4 is not supported. The actual accounting choice by match interaction does not vary depending on the level of estimate uncertainty.[13]

Analysis of Research Questions

Research Question 1 asks whether users' reliability and future profitability judgments will be more extreme when a mismatch occurs between actual and expected accounting choices than when a match occurs. Results in Table 3 reveal that the actual accounting choice by match interaction does not significantly predict participants' reliability assessments or their future profitability assessments. Research Question 2 asks whether assessments of reliability and future profitability mediate the effects of actual accounting choice and match/mismatch on users' investment decisions. Table 3 shows that these assessments do not mediate the investment decision.

Additional Analysis of Credibility Judgments

Mediation analysis reveals that, as predicted in H3, management credibility is the primary mediator of the actual accounting choice by match interaction. This result could be driven by two underlying forces. First, it could be because of the surprise associated with a mismatch, as argued by EVT. Second, it could be because of the information signal contained in management's choice. Below, I consider the effect that each of these forces has on credibility assessments.

The effect of surprise on credibility depends on whether there is a match or a mismatch. In the case of a match, I expect that credibility assessments will either increase or stay the same. As for a mismatch, I expect that credibility assessments will increase if a firm chooses recognition and decrease if the firm chooses disclosure. Therefore, if surprise drives credibility assessments, the result would be a two-way interaction between actual choice and match.

The effect of the information signal on credibility depends on the accounting choice and the level of estimate uncertainty. If estimate uncertainty is high, then recognition will decrease credibility assessments and disclosure will increase credibility assessments. Conversely, if estimate uncertainty is low, then recognition will increase credibility assessments and disclosure will decrease credibility assessments. Therefore, if the information signal drives credibility assessments, a two-way interaction between actual choice and estimate uncertainty would result. Finally, if both surprise *and* the information signal drive credibility assessments a three-way interaction among actual choice, match, and estimate uncertainty would result.

---

[13] I also examined the amount of time that participants took to complete the case. I assume that participants who receive a mismatch spend more time than do participants who receive a match. A t-test comparing the average number of minutes spent to complete the case reveals that participants who received a mismatch took marginally longer than did participants who received a match ($t = 1.34$, $p = 0.09$).

Participants' average credibility assessments are reported in Table 1, Panel A. Table 1, Panel B shows that the two-way interaction between actual choice and match is significant (F = 31.49, p = 0.00). However, the two-way interaction between actual choice and estimate uncertainty is not significant (F = 0.00, p > 0.50). In addition, the three-way interaction among actual choice, match, and estimate uncertainty is not significant (F = 0.09, p > 0.50). Therefore, participants' credibility assessments likely are driven by the surprise associated with a mismatch and not by the information signal contained in management's choice. In Experiment 2, I provide additional evidence consistent with this conclusion.

Discussion

The evidence from Experiment 1 indicates that users' reactions to firms' accounting choices depend partially on their expectations about accounting choice. I find that users' reactions appear to be determined primarily by their perceptions of management's credibility. Specifically, when a firm chooses to recognize stock-based compensation, users invest more in the case of a mismatch than a match. Similarly, when the firm chooses to disclose stock-based compensation, users invest less as the result of a mismatch than a match.

IV. EXPERIMENT 2

In Experiment 1, I predict and find stronger reactions to a mismatch than to a match. As discussed in the additional analysis above, these effects could be driven by (1) the surprise associated with a mismatch and/or (2) the information signal associated with the firm's accounting choice. Therefore, the first purpose of Experiment 2 is to disentangle these two explanations by using a control condition that holds constant the information signal but not the surprise associated with a mismatch.

In Experiment 1, I expected that participants who received a mismatch would be more likely to search for and acquire information about estimate uncertainty. Contrary to my expectations, I did not find a three-way interaction among actual accounting choice, match, and estimate uncertainty. The main effect of estimate uncertainty is consistent with all participants in Experiment 1 viewing the estimate uncertainty information regardless of whether they experienced a match or a mismatch. Therefore, the second purpose of Experiment 2 is to use a different method and reexamine users' propensity to search for additional information.

The third purpose of Experiment 2 is to demonstrate the generality of the effects obtained in Experiment 1. First, to demonstrate that expectations formed for a variety of reasons can affect financial statement users, I set expectations in terms of the firm's past choices instead of the industry norm. Second, I extend the accounting choice setting beyond recognition/disclosure to a context in which preparers have discretion over accounting estimates. Specifically, in Experiment 2, I examine a setting in which the firm has elected to extend the useful life associated with its fixed assets.

Method

Participants

Eighty-two students enrolled in a graduate-level financial statement analysis (FSA) course at a large midwestern university were asked via email to participate in the experiment in exchange for course credit. A total of 69 students (84 percent) participated during April 2008. Participants had taken an average of seven accounting classes and were an average of 23 years old. Sixty-two percent were either M.B.A. or master's of accountancy students.

Ninety-one percent had previously evaluated a firm's performance by analyzing its financial statements. Seventy-two percent are male.[14]

Design

Participants are randomly assigned to one of three treatments in a 1 × 3 between-subjects design that varies whether the firm's accounting choice matches participants' expectations (match versus mismatch versus control). Participants in the control condition receive the same accounting choice and, hence, information signal as participants in the mismatch condition, but do not have expectations about the firm's accounting choice.

The design enables three specific comparisons that can shed light on the results from Experiment 1. First, comparing the match and mismatch conditions replicates the results from Experiment 1. Second, comparing the match and control conditions tests whether there is an effect associated with the information signal contained in management's accounting choice. Finally, comparing the mismatch and control conditions tests whether there is an effect of expectations beyond the effect of the information signal.

Materials and Procedure

The experiment is administered via a three-part web-based instrument. In part 1, participants in the match and mismatch conditions are provided with the property, plant, and equipment footnote disclosure from the 2006 annual report of XYZ Co. that establishes their expectations. Participants in the match (mismatch) condition are told that the firm's only fixed asset, its office building, currently has a 40- (30-) year useful life. Therefore, participants in the match (mismatch) condition are told to expect the firm to use a 40- (30-) year life in the future. Participants in the control condition received no information from the firm's 2006 annual report, and, accordingly, do not have any expectations set about the useful life of the firm's fixed asset. Participants in the match and mismatch conditions then respond to a comprehension-check question and make preliminary assessments of XYZ's management credibility, net income reliability, and future profitability.

Part 2 presents information from XYZ's 2007 annual report. In the match condition, this information consists of the property, plant, and equipment footnote, which indicates a 40-year useful life. In the mismatch and control conditions, this information consists of the property, plant, and equipment footnote, which indicates a 40-year useful life, and the change in accounting estimate footnote, which explains that the useful life has been extended from 30 years to 40 years and that the result of this change is a $.10 increase in EPS. The text of all footnotes provided to participants is based on examples from *Intermediate Accounting,* 12th Edition (Kieso et al. 2007) and is shown in the Appendix.

After reviewing this information, participants in the match and mismatch conditions respond to manipulation checks. Participants in the control condition answer a question that asks them to state the useful life used in 2007. Then all participants answer two questions that examine how likely they are to seek additional fixed-asset related information about XYZ. All participants also rate XYZ's management credibility, net income reliability, and future profitability. Participants conclude the experiment by answering demographic questions.

---

[14] Additional demographic information collected reveals that 87 percent considered English their first language. Participants had an average of one year of work experience and had taken an average of 2.7 finance courses. Sixty-two percent were accounting majors, 26 percent were finance or security analysis majors, and 12 percent were another major. As was the case in Experiment 1, inclusion or exclusion of the demographic variables does not significantly affect the results.

Dependent Variables

Consistent with Experiment 1, the measures elicited in Experiment 2 include assessments of management credibility, net income reliability, and future profitability. Although credibility was the only significant judgment in Experiment 1, I elected to elicit reliability and future profitability in Experiment 2 because much accounting research has argued for the importance of these variables. Thus, it is important to verify that the pattern of results observed in Experiment 1 hold in Experiment 2. Experiment 2 also measures participants' propensity to seek additional information specific to XYZ's fixed assets.[15]

Unlike Experiment 1, which assessed management credibility using a single question, Experiment 2 assesses management credibility using the six-item scale shown in the Appendix. The items are based on prior research by Mercer (2005) and Tan et al. (2002). These questions have a Cronbach's alpha of 0.83, which indicates that they capture a common underlying construct. Therefore, I use the average response to these six questions in my analysis. Perceived reliability and future profitability are measured using the same scales used in Experiment 1.

I measure participants' propensity to seek additional information about XYZ's fixed assets in two ways. First, I ask them how likely they would be to look for additional information about XYZ's fixed assets using a nine-point scale, with endpoints labeled "Not at All Likely" and "Very Likely" and a midpoint labeled "Somewhat Likely." Second, I use an approach developed in the psychology literature to measure attributional processing (e.g., Pyszczynski and Greenberg 1981). Specifically I ask participants to select from a list of eight questions the four that they would most like to get answered as part of an overall analysis of XYZ. Four of these questions would give reasons why a firm would extend the useful life of a fixed asset, and the remaining four questions provide other information that would be of interest to FSA students, such as the capital structure of the firm. All questions were derived based on guidance from *Business Analysis & Valuation* (Palepu et al. 2000). All questions are listed in the Appendix. The software program randomized the order in which these questions were presented.

Results

Manipulation Checks

Responses to the manipulation-check questions indicate that the manipulation was successful. Among participants in the match (mismatch) condition, 88 percent (83 percent) correctly answered that XYZ's actual choice was consistent (inconsistent) with expectations. Excluding the participants who failed this manipulation check does not affect the results. Therefore, analyses include these participants.

As in Experiment 1, the manipulation in Experiment 2 is deemed successful if participants viewed a mismatch between actual and expected accounting choice as less expected and more surprising than a match. Participants rated XYZ's choice as significantly less expected ($t = 4.09$, $p = 0.00$) if they received a mismatch (mean $= 5.79$) than if they received a match (mean $= 3.46$). Similarly, participants rated XYZ's choice as significantly more surprising ($t = 3.35$, $p = 0.00$) if they received a mismatch (mean = 4.88) than if they received a match (mean = 2.88).

---

[15] Participants in Experiment 2 received extra credit in exchange for participation. The amount of extra credit offered dictated that the time required for the task be limited. Whereas participants in Experiment 1 received considerable information from which to make a meaningful investment decision (i.e., P/E information, industry information, information about a competitor, and pricing information for the firm), participants in Experiment 2 do not. Therefore, participants in Experiment 2 are not asked to provide an investment decision. This choice simplifies and shortens participants' task considerably. However, it also precludes me from testing H2, H3, and RQ2 in Experiment 2.

Participants in the control condition are asked to indicate the useful life that XYZ used in 2007. All control condition participants correctly answer that XYZ's fixed asset had a 40-year useful life in 2007. Therefore, control condition participants appear to have attended to the information provided to them.[16]

Hypothesis Tests

In Experiment 2, increasing the useful life of the firm's fixed asset serves to decrease the firm's depreciation expense and constitutes a negative expectancy violation. Therefore, support for H1 and RQ1 is evidenced by users reacting less favorably to management's choice of a useful life for its fixed assets if that choice does not match users' expectations. Based on the results from Experiment 1, I expect a significant effect for credibility (H1) but not for reliability or future profitability (RQ1). Planned contrasts reported in Table 4, Panel B, reveal that participants rated XYZ's management as significantly less credible ($t = 1.71$, $p = 0.05$) if they received a mismatch (mean = 4.78) than if they received a match (mean = 5.31).[17] In addition, participants rated XYZ's net income as significantly less reliable ($t = 1.75$, $p = 0.04$) if they received a mismatch (mean = 4.13) than if they received a match (mean = 4.92). Participants do not rate XYZ's future profitability as significantly different ($t = 1.24$, $p = 0.11$) in the mismatch condition (mean = 4.61) relative to the match condition (mean = 5.13).

As discussed above, these results could be because either the accounting choice was unexpected or because of the information signal associated with management's choice. Therefore, it is important to compare the mismatch condition to the control condition to parse out the effect of participants' expectations. As shown in Table 4, Panel B, a planned contrast reveals that participants rated XYZ's management as significantly less credible ($t = 1.79$, $p = 0.04$) if they received a mismatch than if they were in the control condition. However, the reliability and future profitability ratings do not significantly differ between these two conditions. Furthermore, the credibility, reliability, and future profitability ratings do not statistically differ in either the match or control conditions. Therefore, the information signal embedded in management's accounting choice does not appear to account for the significant difference between the credibility assessments in the match and mismatch conditions. Instead, this difference is driven by the fact that management's choice was inconsistent with participants' expectations supporting H1.[18]

Hypothesis 4 predicts that participants who receive a mismatch are more likely to seek an explanation for the firm's choice than participants who receive a match. As shown in Table 4, Panel B, planned contrasts reveal that participants are significantly more likely to seek additional information ($t = 2.10$, $p = 0.02$) in the mismatch condition (mean = 7.33) than in the match condition (mean = 5.96).

---

[16] Participants in the control condition were not asked whether XYZ's choice was consistent or inconsistent with their expectations because the control condition was intentionally absent expectations. Similarly, participants in the control condition were not asked to rate the unexpectedness and surprise associated with XYZ's choice. Asking questions about participants' expectations in the control condition would have served to highlight expectations, thus weakening the effectiveness of the control condition.

[17] Furthermore, a planned contrast reveals that the change in credibility judgments from pre to post is significantly greater ($t = 2.25$, $p = 0.03$) in the mismatch condition (-0.75) than in the match condition (-0.16).

[18] Table 4, Panel B also reveals there is no significant difference in the credibility, future profitability, and reliability prejudgments, indicating that the reported results are driven by the interplay between actual and expected accounting choices.

TABLE 4
Experiment 2 Results

Panel A: Cell Means[a]

	Match n = 24	Control n = 21	Mismatch n = 24
Credibility Prejudgment	5.47	NA	5.53
Reliability Prejudgment	4.96	NA	4.88
Future Profitability Prejudgment	5.21	NA	5.08
Seek Additional Info	5.96	6.19	7.33
Credibility Post Judgment	5.31	5.36	4.78
Reliability Post Judgment	4.92	4.52	4.13
Future Profitability Post Judgment	5.13	4.92	4.61

Panel B: Planned Contrasts: t-value (p-value)[b]

	Match versus Mismatch	Mismatch versus Control	Match versus Control
Credibility Prejudgment	0.19 (0.43)	NA	NA
Reliability Prejudgment	0.17 (0.44)	NA	NA
Future Profitability Prejudgment	0.40 (0.35)	NA	NA
Seek Additional Info	**2.10*** **(0.02)**	**1.69** **(0.05)**	0.34 (0.37)
Credibility Post Judgment	**1.71** **(0.05)**	**1.79** **(0.04)**	0.14 (0.45)
Reliability Post Judgment	**1.75** **(0.04)**	0.85 (0.20)	0.84 (0.20)
Future Profitability Post Judgment	1.24 (0.11)	0.67 (0.25)	0.53 (0.30)

---

* Significant differences between treatments are bold.

[a] Panel A provides the mean ratings for all dependent measures elicited from participants in Experiment 2. All prejudgments are elicited after setting participants' expectations but before providing them with the firm's actual accounting choice. All post judgments are elicited after providing participants with the firm's actual choice. The credibility pre- and postjudgments are the average of participants' responses to the six questions shown in the Appendix. The reliability pre- and postjudgments are elicited on a nine-point scale with the endpoints of the scale labeled "Very unreliable" and "Very reliable" and a midpoint labeled "Somewhat reliable." The future profitability pre- and postjudgments are elicited on a nine-point scale with the endpoints of the scale labeled "Very unfavorable" and "Very favorable" and the midpoint labeled "Neutral." Participants' propensity to seek additional information is elicited on a nine-point scale with endpoints labeled "Not at all likely" and "Very likely" and a midpoint labeled "Somewhat likely."

[b] Panel B provides the t-values and one-sided p-values associated with all planned comparisons.

I also examine participants' propensity to search for additional information about XYZ's fixed assets by asking participants to select four out of eight questions that they would like answered as part of an overall analysis of XYZ. Four of these items are directly related to XYZ's fixed assets. The other four are related to XYZ in general. Therefore, participants could select as many as four and as few as zero questions related to XYZ's fixed assets.

Table 5 shows the number of participants in each experimental condition who fall into each category. A Fisher's exact test reveals that the distribution of participants across all five categories in the mismatch condition is significantly different than in the match condition ($p = 0.10$). It is important to note that this evidence is suggestive, rather than conclusive, because alternative ways of interpreting these distributions lead to alternative inferences. However, when coupled with the scale measure reported above, these results provide at least some preliminary evidence that participants receiving a mismatch are more likely to seek additional information about the reason for the mismatch than are participants who receive a match. Therefore, the results partially support H4.

V. GENERAL DISCUSSION

This research investigates how users respond to a firm's discretionary accounting choices when those choice either match or do not match users' expectations. Based on EVT, I predict and find that users react more strongly to a mismatch than to a match. Taken together, results from Experiments 1 and 2 demonstrate that these effects hold in a variety of accounting settings and are driven by users' expectations rather than by the information signal in the firm's accounting choice. Results from Experiment 1 also reveal that the interaction between actual accounting choice and the extent to which it matches expectations affects assessments of management's credibility, which then affect investment decisions. Also consistent with EVT, I find some evidence, albeit mixed, that when a mismatch between actual and expected accounting choices occurs, users are likelier than when there is a match to seek more information that can explain the firm's choice.

My findings are relevant to accounting researchers, regulators, standard-setters, and managers because they demonstrate the importance of considering users' expectations when predicting how users will react to a firm's accounting choices. In my settings, I find that these reactions are primarily driven by perceptions of management credibility. Prior research has found that in some circumstances credibility effects can be fleeting (Mercer 2005), thereby implying that credibility effects are not a significant determinant of behavior. However, in Experiment 1, I find that credibility effects can drive investment decisions, thereby

TABLE 5
Experiment 2 Results
Distribution of Participants Seeking Additional Information about XYZ's Fixed Assets

Number of Questions Selected that Relate to XYZ's Fixed Assets	Match	Control	Mismatch
4	10	15	8
3	6	3	12
2	3	2	4
1	4	1	0
0	1	0	0

The table above groups the participants in the experimental conditions based on the number of fixed-asset related questions they selected. Participants' were provided with a list of eight questions and were asked to select the four questions that they would most like answered as part of an overall analysis of XYZ. Four of these questions relate to reasons why a firm would extend the useful life of a fixed asset. The remaining four questions provide other information about the firm. All questions are listed in the Appendix. Participants could select as many as four and as few as zero questions related to XYZ's fixed assets.

resulting in longer-term ramifications if users fail to adjust their investment decisions after credibility effects have subsided. Furthermore, I find some evidence that when expectations about an accounting choice are inconsistent with management's actual choice, this inconsistency affects users' propensity to search for additional information. This finding has implications for researchers, standard-setters, and regulators concerned with the factors that affect consumption of additional disclosures.

My results are subject to a number of limitations that suggest avenues for future research. In Experiment 1, I find that users' perceptions of management's credibility drive investment decisions. However, credibility assessments certainly are not the only drivers of investment decisions in all contexts. Future research could examine additional factors that determine when credibility assessments do and do not mediate investment decisions. Furthermore, in Experiment 2, I do not find a significant difference between the match and the control conditions. This suggests that users are not responding to the signal contained in a firm's choice to extend the useful life of its fixed assets. This finding seems surprising, given prior literature on signaling in accounting. Future research could reconcile these results. In Experiment 2, I ask participants to rate how likely they are to search for additional information. However, I do not require them to engage in an actual search. Thus, it is an open question as to whether participants' stated intentions would result in actual search behavior. Finally, I examine two settings in which cost recognition is preferred to disclosure. There may be other settings in which users would prefer cost disclosure to recognition.

APPENDIX
EXCERPTS FROM EXPERIMENTAL MATERIALS

Manipulation of Estimate Uncertainty in Experiment 1

Low Estimate Uncertainty Treatment

Assumptions Underlying Stock-based Compensation Expense: The fair value of stock-based compensation is calculated using the Black-Scholes model. CAP's options *are* very similar to those that the Black-Scholes model was designed to value and CAP *has* a long history of accurately estimating the assumptions of the Black-Scholes model. Therefore, **the firm *is* very confident in the accuracy of its estimate of the fair value of stock-based compensation.**

High Estimate Uncertainty Treatment

Assumptions Underlying Stock-based Compensation Expense: The fair value of stock-based compensation is calculated using the Black-Scholes model. CAP's options *are not* very similar to those that the Black-Scholes model was designed to value and CAP *does not have* a long history of accurately estimating the assumptions of the Black-Scholes model. Therefore, **the firm *is not* very confident in the accuracy of its estimate of the fair value of stock-based compensation.**

Note: Participants received one of the two footnotes shown above as dictated by their treatment condition. Italicized text emphasizes differences between the two treatments for the purposes of this figure only. All text presented to participants was nonitalicized. All bolded and nonbolded text above appears exactly as it was presented to participants.

EPS Information Provided to Participants in Experiment 1

Panel A: Actual Recognition, Mismatch Treatments

	12/31/2005		12/31/2004	
	Reported on Income Statement (Reduced by Stock Options)	**Pro Forma in Footnotes** (Not Reduced by Stock Options)	**Reported on Income Statement** (Reduced by Stock Options)	**Pro Forma in Footnotes** (Not Reduced by Stock Options)
CAP's EPS	2.00	2.09	1.93	2.02

	12/31/2005		12/31/2004	
	Reported on Income Statement (Not Reduced by Stock Options)	**Pro Forma in Footnotes** (Reduced by Stock Options)	**Reported on Income Statement** (Not Reduced by Stock Options)	**Pro Forma in Footnotes** (Reduced by Stock Options)
Parts Inc.'s EPS	2.07	2.00	2.00	1.93

Panel B: Actual Disclosure, Mismatch Treatments

	12/31/2005		12/31/2004	
	Reported on Income Statement (Not Reduced by Stock Options)	**Pro Forma in Footnotes** (Reduced by Stock Options)	**Reported on Income Statement** (Not Reduced by Stock Options)	**Pro Forma in Footnotes** (Reduced by Stock Options)
CAP's EPS	2.09	2.00	2.02	1.93

	12/31/2005		12/31/2004	
	Reported on Income Statement (Reduced by Stock Options)	**Pro Forma in Footnotes** (Not Reduced by Stock Options)	**Reported on Income Statement** (Reduced by Stock Options)	**Pro Forma in Footnotes** (Not Reduced by Stock Options)
Parts Inc.'s EPS	2.00	2.07	1.93	2.00

Note: The table provided in Panel A above is taken from the Actual Recognition, Mismatch treatments. In the Match treatments, the labels "Reduced by Stock Options" and "Not Reduced by Stock Options" and the corresponding EPS amounts are reversed for Parts, Inc. In the Actual Disclosure treatments, the labels "Reduced by Stock Options" and "Not Reduced by Stock Options" and the corresponding EPS amounts are reversed for CAP, as shown in Panel B for the Actual Disclosure, Mismatch treatments.

Footnotes Provided in Experiment 2

Information Provided in Part 1 (from 2006 annual report)

Match Condition

Note 2. Property, Plant, and Equipment

Property, plant, and equipment are recorded at cost reduced by accumulated depreciation. Depreciation expense is recognized over the asset's estimated useful life using the straight-line method. The company's only depreciable asset is the building that houses its corporate headquarters. The building is depreciated over an estimated useful life of 40 years.

Mismatch Condition

Note 2. Property, Plant, and Equipment

Property, plant, and equipment are recorded at cost reduced by accumulated depreciation. Depreciation expense is recognized over the asset's estimated useful life using the straight-line method. The company's only depreciable asset is the building that houses its corporate headquarters. The building is depreciated over an estimated useful life of 30 years.

Control Condition

Part 1 not provided.

Information Provided in Part 2 (from 2007 annual report)

Match Condition

Note 2. Property, Plant, and Equipment

Property, plant, and equipment are recorded at cost reduced by accumulated depreciation. Depreciation expense is recognized over the asset's estimated useful life using the straight-line method. The company's only depreciable asset is the building that houses its corporate headquarters. The building is depreciated over an estimated useful life of 40 years.

Mismatch and Control Conditions

Note 2. Property, Plant, and Equipment

Property, plant, and equipment are recorded at cost reduced by accumulated depreciation. Depreciation expense is recognized over the asset's estimated useful life using the straight-line method. The company's only depreciable asset is the building that houses its corporate headquarters. The building is depreciated over an estimated useful life of 40 years.

Note 11. Change in Accounting Estimate

In the current year, the company revised its estimate of the useful life of the building that houses its headquarters. This is the company's only depreciable asset and previously had a useful life of 30 years. The useful life is now extended to 40 years. This change has the effect of reducing depreciation expense and increasing net income by approximately $991,000 ($.10 per share).

Note: Only participants in the match and mismatch conditions received Part 1, which established their expectations. Participants in all conditions received Part 2. In the match condition, the firm's useful life for its fixed asset did not change from 2006 to 2007. In the mismatch and control conditions, the useful life increased from 30 years to 40 years.

Management Credibility Questions in Experiment 2

When it comes to providing financial disclosures, I believe XYZ's management ...

	Strongly Disagree	2	3	4	Neither Agree nor Disagree	6	7	8	Strongly Agree
is knowledgeable.	○	○	○	○	○	○	○	○	○
may **not** be competent.	○	○	○	○	○	○	○	○	○
is qualified.	○	○	○	○	○	○	○	○	○
is trustworthy.	○	○	○	○	○	○	○	○	○
is honest.	○	○	○	○	○	○	○	○	○
may **not** be forthcoming.	○	○	○	○	○	○	○	○	○

Note: Participants evaluated each statement twice. The first evaluation occurred after they received the information that set their expectations but before the information that revealed XYZ's actual choice. The second evaluation occurred after they received the information about XYZ's actual choice. Items 2 and 6 were reverse-scored. The credibility measure used in the analysis is a composite measure formed by averaging each participant's response to the six questions.

Attributional Processing Questions in Experiment 2

If, as part of an overall analysis of XYZ, you could receive answers to only 4 of the 8 questions listed below, which 4 questions would you choose?

Select at least 4 responses and no more than 4 responses.

- How long does XYZ expect its building to remain in service?
- What is the useful life on fixed assets for other firms in this industry?
- Has XYZ made any recent improvements to its headquarters?
- What is XYZ's capital structure?
- What is the useful life on fixed assets for other firms in this party of the country?
- What is XYZ's credit policy?
- How well does XYZ manage its inventory?
- Who are XYZ's primary customers?

Note: The software program randomized the order in which the eight questions were presented. Questions 1, 2, 3, and 5 are specific to XYZ's fixed assets. The others are general questions that participants would want answered if they were analyzing a firm's financial statements.

REFERENCES

Aboody, D., M. Barth, and R. Kasznik. 2004. Firms' voluntary recognition of stock-based compensation expense. *Journal of Accounting Research* 42: 123–150.

Accounting Principles Board (APB). 1972. *Accounting for Stock Issues to Employees.* APB Opinion No. 25. New York, NY: APB.

Ackert, L. F., B. K. Church, and M. Shehata. 1996. What affects individuals' decisions to acquire forecasted information? *Contemporary Accounting Research* 13 (Fall): 379–399.

———, ———, and ———. 1997. An experimental examination of the effects of forecast bias on individuals' use of forecasted information. *Journal of Accounting Research* 35 (1): 25–42.

Allessandri, T., and R. Khan. 2006. Market performance and deviance from industry norms: (Mis)alignment of organizational risk and industry risk. *Journal of Business Research* 59 (10/11): 1105–1115.

Baron, R. M., and D. A. Kenny. 1986. The moderator-mediator variable distinction in social psychological research: Conceptual, strategic, and statistical considerations. *Journal of Personality and Social Psychology* 51 (6): 1173–1182.

Bartov, E., D. Givoly, and C. Hayn. 2002. The rewards to meeting or beating earnings expectations. *Journal of Accounting and Economics* 33 (2): 173–204.

Bevan, J. L. 2003. Expectancy violation theory and sexual resistance in close, cross-sex relationships. *Communication Monographs* 70 (1): 68–82.

Burgoon, J. K. 1983. Nonverbal violations of expectations. In *Nonverbal Interaction*, edited by J. M. Wiemann, and R. P. Harrison, 77–112. Beverly Hills, CA: Sage.

———, and J. L. Hale. 1988. Nonverbal expectancy violations: Model elaboration and application to immediacy behaviors. *Communication Monographs* 55: 58–79.

———, D. A. Newton, J. B. Walther, and E. J. Baesler. 1989. Nonverbal expectancy violations and conversational involvement. *Journal of Nonverbal Behavior* 13: 97–120.

———, and B. A. LePoire. 1993. Effects of communication expectancies, actual communication, and expectancy disconfirmation on evaluations of communicators and their communication behavior. *Human Communication Research* 20: 75–107.

———, and M. Burgoon. 2001. Expectancy Theories. In *The New Handbook of Language and Social Psychology*, edited by W. P. Robinson, and H. Giles, 79–99. London, U.K.: John Wiley & Sons, Ltd.

Campo, S., K. A. Cameron, D. Brossard, and M. S. Frazer. 2004. Social norms and expectancy violation theories: Assessing the effectiveness of health communication campaigns. *Communication Monographs* 71 (4): 448–470.

———, and K. A. Cameron. 2006. Differential effects of exposure to social norms campaigns: A cause for concern. *Health Communication* 19 (3): 209–219.

Chatterjee, S., M. Lubatkin, and W. Schulze. 1999. Toward a strategic theory of risk premium: Moving beyond CAPM. *Academy of Management Review* 24 (3): 556–567.

Financial Accounting Standards Board (FASB). 1984. *Recognition and Measurement in Financial Statements of Business Enterprises.* Statement of Financial Accounting Concepts No. 5. Stamford, CT: FASB.

———. 1995. *Accounting for Stock-Based Compensation.* Statement of Financial Accounting Standards No. 123. Norwalk, CT: FASB.

———. 2004. *Shared-Based Payments.* Statement of Financial Accounting Standards No. 123R. Norwalk, CT: FASB.

Frederickson, J. R., F. Hodge, and J. Pratt. 2006. The evolution of stock option accounting: Disclosure, voluntary recognition, mandated recognition, and management disavowals. *The Accounting Review* 81 (5): 1073–1093.

Griffin, K. 1967. The contribution of studies of source credibility to a theory of interpersonal trust in the communication process. *Psychological Bulletin* 68 (2): 104–120.

Hirst, D. E., L. Koonce, and P. J. Simko. 1995. Investor reactions to financial analysts' research reports. *Journal of Accounting Research* 33 (2): 335–351.

———, and P. Hopkins. 1998. Comprehensive income reporting and analysts' valuation judgments. *Journal of Accounting Research* 36 (Supplement): 47–75.

Hodge, F., S. Kennedy, and L. Maines. 2004. Does search-facilitating technology improve the transparency of financial reporting? *The Accounting Review* 79 (3): 687–703.

———, P. Hopkins, and J. Pratt. 2006. Management reporting incentives and classification credibility: The effects of reporting discretion and reputation. *Accounting, Organizations and Society* 31 (7): 623–634.

Hopkins, P. E. 1996. The effect of financial statement classification of hybrid financial instruments on financial analysts' stock price judgments. *Journal of Accounting Research* 34 (Supplement): 33–50.

———, R. W. Houston, and M. F. Peters. 2000. Purchase, pooling, and equity analysts' valuation judgments. *The Accounting Review* 75 (3): 257–281.

Houser, M. L. 2005. Are we violating their expectations? Instructor communication expectations of traditional and nontraditional students. *Communication Quarterly* 53 (2): 213–228.

Hovland, C., I. Janis, and H. Kelley. 1953. *Communication and Persuasion.* New Haven, CT: Yale University Press.

Kenny, D. A., D. A. Kashy, and N. Bolger. 1998. Data analysis in social psychology. In *The Handbook of Social Psychology*, edited by D. T. Gilbert, and S. T. Fiske., Vol. 1, 233–265. 4th edition. New York, NY: McGraw-Hill.

Kieso, D. E., J. J. Weygandt, and T. D. Warfield. 2007. *Intermediate Accounting.* 12th edition. Hoboken, NJ: John Wiley & Sons.

Kleinbaum, D. G., L. L. Kupper, K. E. Muller, and A. Nizam. 1998. *Applied Regression Analysis and Multivariable Methods.* 3rd edition. Pacific Grove, CA: Brooks/Cole Publishing Company.

LePoire, B. A., and J. K. Burgoon. 1994. Two contrasting explanations of involvement violations: Expectancy violations theory and discrepancy arousal theory. *Human Communication Research* 20: 560–591.

Maines, L. A., L. S. McDaniel, and M. S. Harris. 1997. Implications of proposed segment reporting standards for financial analysts' investment judgments. *Journal of Accounting Research* 35 (Supplement): 1–24.

———, and ———. 2000. Effects of comprehensive income volatility on non-professional investors' judgments: the role of presentation format. *The Accounting Review* 75 (April): 179–207.

Mercer, M. 2005. The fleeting effects of disclosure forthcomingness on management's reporting credibility. *The Accounting Review* 80 (2): 723–744.

Nelson, M. W., S. D. Smith, and Z. Palmrose. 2005. The effect of quantitative materiality approach on auditors' adjustment decisions. *The Accounting Review* 80 (3): 897–920.

Palepu, K. G., P. M. Healy, and V. Bernard. 2000. *Business Analysis & Valuation: Using Financial Statements.* 2nd edition. Cincinnati, OH: South-Western College Publishing.

Pyszczynski, T. A., and J. Greenberg. 1981. Role of disconfirmed expectancies in the instigation of attributional processing. *Journal of Personality and Social Psychology* 40 (1): 31–38.

Tan, H. T., R. Libby, and J. E. Hunton. 2002. Analysts' reactions to earnings preannouncement strategies. *Journal of Accounting Research* 40 (1): 223–246.

THE ACCOUNTING REVIEW
Vol. 84, No. 5
2009
pp. 1495–1520

American Accounting Association
DOI: 10.2308/accr.2009.84.5.1495

Stealth Disclosure of Accounting Restatements

Rebecca Files
The University of Texas at Dallas

Edward P. Swanson
Senyo Tse
Texas A&M University

ABSTRACT: Managers exercise considerable discretion over how they announce an accounting restatement in a press release. Some firms issue a press release that discloses the restatement in the headline (high prominence). Others provide a press release with a headline on a different subject (for example, earnings news) but describe the restatement in the body of the release (medium prominence). The remaining firms discuss the restatement at the end of the press release in a footnote to operating results (low prominence). Mean three-day returns differ considerably across these three categories of prominence (−8.3, −4.0, and −1.5 percent, respectively). We find that disclosure prominence is significantly negatively associated with returns in a model that controls for the seriousness of the GAAP violation, restatement magnitude, other restatement characteristics, and potential endogeneity. Similarly, we find the likelihood of class action lawsuits is significantly reduced with less prominent disclosure.

Keywords: *accounting restatements; press release disclosure; voluntary disclosure.*

Data Availability: *The data used in this study are publicly available from the sources indicated in the text.*

I. INTRODUCTION

Press releases issued by companies identified in a study of accounting restatements prepared for Congress (GAO 2002) reveal that many corporate managers provide only minimal disclosure when announcing a restatement. We observe three levels of disclosure. Some companies disclose their restatement prominently in the headline of a

---

We thank Steve Kachelmeier (editor), Steve Huddart, and two anonymous referees. We also thank Nerissa Brown, Eric Hirst, Leslie Davis Hodder, Michael Kimbrough, Robert Libby, Tom Omer, Devin Shanthikumar, Nate Sharp, and workshop participants at Texas A&M University for their suggestions. We also benefited from comments at the 2007 AAA Annual Meeting, 2007 CFEA Conference at NYU, and the 2008 AAA FARS Conference. We are indebted to Jason Call, Vikas Hegde, Harrison Liu, Rishekesy Gorantala, and Natasha Thomas for their research assistance. The authors gratefully acknowledge financial support from the Mays Business School of Texas A&M. Professor Swanson also acknowledges funding from the Durst Chair in Accounting.

Editor's note: Accepted by Steven Kachelmeier, with thanks to Steve Huddart for serving as editor on a previous version.

Submitted: August 2007
Accepted: November 2008
Published Online: September 2009

press release, usually one that is dedicated to the accounting misstatement (high prominence). Other firms provide less prominent disclosure, typically citing an earnings release in the headline, but still discussing the misstatement in the body of the press release (medium prominence). Most of the remaining firms simply restate prior-period comparative balances in an earnings release, with a footnote briefly explaining that the financial figures for the prior year have been changed (low prominence).

We investigate whether companies providing medium or low prominence disclosure of their restatement benefit from a less negative market reaction and/or a reduced likelihood of litigation. This benefit could occur if investors view a misstatement that is not mentioned in a headline as an unimportant outcome of a company's good-faith effort to implement complex accounting rules or if they miss the disclosure entirely.[1] This issue is closely related to the broader question of whether firms are rewarded or penalized for openness in disclosing bad news, a question that has been studied in the context of earnings warnings. Kasznik and Lev (1995) find that returns for warning firms are lower than returns for similar non-warning firms. Tucker (2007) confirms this result, but reports that warning firms have less favorable non-earnings news (e.g., product news, alliances, and management changes) than no-warning firms, suggesting that warning firms' long-window returns would have been lower than non-warning firms' returns even without the warning. She concludes that firms are *not* penalized in the long-run for providing voluntary management forecasts. Prior studies also consider the rewards and penalties for openness in disclosing bad news in the context of shareholder litigation. Skinner (1994, 1997) conjectures that some firms warn investors in advance of a large negative earnings surprise to avoid litigation. Francis et al. (1994) find, however, that warning companies are more likely to be sued than non-warning firms. Field et al. (2005) control for the endogenous relation between earnings warnings and litigation and conclude that warnings may indeed deter litigation. Existing evidence is therefore mixed about whether firms are rewarded or penalized for openness in disclosing bad news.

Our study provides new evidence on this question, but in the context of accounting misstatements rather than earnings shortfalls.[2] We also examine disclosure *prominence* rather than the presence or absence of disclosure. That is, all firms in our sample disclose restatements in press releases, but do so with varying degrees of prominence. Our first finding is that the magnitude of the market response to a restatement announcement is related to press release format. Three-day returns differ substantially across the three categories of disclosure prominence, averaging −8.3 percent, −4.0 percent, and −1.5 percent for high, medium, and low prominence, respectively. Returns for the high prominence group are statistically different from those for the medium and low prominence groups. Next, we extend the return window to 20 days after the announcement to investigate post-announcement responses to restatements. We find returns of −7.9 percent, −6.4 percent, and −3.2 percent for the high, medium, and low prominence firms, respectively. These returns are considerably less dispersed than the short-window returns, and the 1.5

---

[1] In a 2006 report, Glass Lewis & Co., LLC, a firm that consults on proxy issues and conducts policy-oriented research, makes a similar comment about companies that do not file an amended SEC report or Form 8-K (Reilly 2006a, 2006b). Glass Lewis suggests those companies hope that investors "may dismiss a restatement as relatively minor" if it is tucked "away quietly in the current period's results." The SEC has since explained that restatements "constitute information that is material to investors and thus needs to be formally disclosed in a restatement filing clearly labeled as such" (Reilly 2006a). While the SEC directive is aimed at companies' SEC filings, we focus on the press release because we find little market reaction to the SEC filings.

[2] Since the passage of the Private Securities Litigation Reform Act of 1995, accounting litigation has shifted to firms disclosing the need for an accounting restatement and away from class actions alleging a failure to provide timely warning of an earnings shortfall (Grundfest and Perino 1997).

percent return difference between high and medium prominence is not statistically significant. Apparently, market participants initially underestimate the seriousness of some misstatements disclosed without a headline but subsequently correct their under-reaction.

To determine whether investors' post-restatement responses to medium and low prominence restatements are due to new disclosures, we examine news sources in the 20-day post-announcement period (press releases, SEC filings, news reports, etc.). We find that the average return for firms that have post-restatement news items is not significantly different from zero. In contrast, we find a statistically significant drift of −3.7 percent for firms with no news items in the 20-day period, which suggests that investors further evaluate the original press release information. Analysts appear to play an important role in this evaluation because most of the drift is in companies covered by three or more analysts. These results are consistent with limited attention theory (Hirshleifer and Teoh 2003).

Next, we examine if press release prominence remains negatively associated with market returns when we control for the seriousness of the accounting misstatement (using the magnitude of the restatement, an indicator variable for whether the misstatement is an "error" or an "irregularity," and the level of short selling). For this test, we code prominence as 3, 2, and 1 for high, medium, and low prominence, respectively. We find that this prominence index is highly significant in explaining announcement period returns (−1, +1), but the index is not significantly associated with returns over the longer window (−1, +20). The latter finding is consistent with the negative drift we observe for firms providing medium and low prominence disclosure (discussed above). A correction of the initial mispricing is reassuring, but the initial under-reaction could be economically significant. During the drift period, managers and other investors have an opportunity to sell their stock at a higher price, and investors who purchase shares would (on average) experience a short-term price drop. In addition, a smaller market reaction to the restatement announcement can reduce the likelihood of litigation, as discussed next.

Class action lawyers respond to large price declines that they can link to managers' actions. Prominent misstatement disclosure results in a greater price decline (as discussed above) and makes it easier to link the decline directly to managers' actions. A larger price decline would also increase potential damages if a suit is initiated. We find that the frequency of lawsuits declines monotonically across the three categories of disclosure prominence (27 percent, 16 percent, and 0 percent for the high, medium, and low prominence firms, respectively). The 16 percent litigation rate for medium disclosure suggests that some managers use medium prominence disclosure for an accounting misstatement that plaintiff attorneys view as serious. An interesting question is whether that rate would have been even higher if those firms had provided high prominence (headline) disclosure. We investigate this issue formally by estimating a logistic regression model of the likelihood of litigation. We include controls for misstatement severity, restatement characteristics, other determinants of lawsuits identified in prior research, and announcement-date returns. We find that the prominence index coefficient is positive and significant in the model (even after controlling for endogeneity), suggesting that the likelihood of litigation rises with disclosure prominence. Reducing disclosure prominence by one level (e.g., medium instead of high) reduces the odds of a lawsuit by about half. This result is unchanged when we exclude dismissed lawsuits from the sample, so in contrast to the return differences, the effect seems to be permanent.

An important question follows from this evidence: Why do managers frequently emphasize restatements by disclosing them in a press release headline? One possibility is that restatements occur infrequently, so managers have little experience that would allow them to anticipate how investors and/or litigators will react. Another possibility is that some companies adopt a policy of being forthright in corporate communications, expecting to lower their cost of capital over the long run by reducing uncertainty (Graham et al. 2005).

In Section II, we present the hypotheses. In Section III, we describe the sample. We summarize the multivariate models used to explain returns and litigation in Section IV. In Section V, we present the results based on market returns, including robustness checks and extensions. In Section VI, we report litigation results. We conclude in Section VII.

II. HYPOTHESES

Guidelines about how to write an effective press release stress the importance of the headline. Mahoney (1991, 284) comments, "The headlines drive an idea, instead of simply identifying a subject." Feid (2005, 60) states, "Don't bury the news—a headline should convey the entire key message at the top of the press release. Then the lead paragraph should reiterate it with the basic facts."

To date, the effect of press release format on market reactions has received little attention in the accounting literature. A relevant exception is a study by Bowen et al. (2005), which investigates whether investors' relative response to GAAP earnings versus pro forma earnings depends on their placement in the firm's press release. They find that the relative strength of the price response to each earnings metric is greater if that metric is emphasized by its placement in the press release. Using an approach similar to that in Bowen et al. (2005), we define a three-level index of disclosure prominence. We code as 3 (high prominence) any press release in which the restatement is mentioned in the headline; we use 2 (medium prominence) for any press release in which the restatement is not mentioned in the headline but is discussed in some detail within the body of the report;[3] and we assign 1 (low prominence) to those press releases that only mention the restatement in the footnotes to comparative-period results.[4]

Concurrent research by Gordon et al. (2008) also examines the effect of press release characteristics on returns in a three-day window centered on the restatement announcement. Of direct relevance to our study, they find that firms reporting the restatement in the headline experience a larger price decline than other firms, and this finding remains after they control for restatement characteristics. In other respects, their paper is quite different from our study.[5] Myers et al. (2008) consider a more severe form of stealth restatements: namely, restatements that are never disclosed in a press release. Thus, in contrast to our focus on disclosure prominence within press releases, Myers et al. (2008) examine restating firms that forgo press release (or Form 8-K) disclosure of their restatements. They classify restatements announced in a press release or Form 8-K as the most transparent; followed by those announced in an amended SEC filing; and followed by restatements announced only in a regular, periodic SEC filing with no press release or 8-K filing. They find less negative market reactions to restatements disclosed only in routine SEC filings. Myers et al. (2008) also explore whether firms systematically bundle restatement announcements with earnings news and examine the determinants of firms' restatement disclosure choices.

---

[3] These firms disclose a wide variety of company information in their press releases, although most of them (58 percent) report current earnings information. Other types of information include a new credit facility (Sybron Dental 2002), settlement of past litigation charges (Cognos 2002), and adoption of new tax rules (First American Financial Corporation 2000). Kothari et al. (2008, 2) posit that managers often delay disclosing bad news, apparently gambling that subsequent corporate events (i.e., future corporate turnaround or restructuring) will allow them to "bury" the bad news.

[4] Bowen et al. (2005) use a four-point scale for each earnings metric: 1 = discussed in the headline; 2 = discussed in the 1st or 2nd paragraph; 3 = discussed further down in the body of the press release; and 4 = only discussed in the financial statements provided at the end of the release (i.e., the footnotes). Our index combines scores 2 and 3 into one measure.

[5] Gordon et al. (2008) focus on determining whether disclosure credibility is associated with market reactions to restatement announcements. Disclosure credibility consists of situational incentives at the time of the disclosure (which they argue is homogenous among restating firms), management credibility (determined by the amount and tone of prior discretionary disclosures), and the characteristics of the restatement announcement. Their research therefore complements and extends the capital markets analysis in our paper. Note that Gordon et al. (2008) do not consider whether restatement disclosure practices influence the likelihood of litigation.

Theory differs about whether the format of a press release should affect investor response. In an efficient market, security prices respond promptly and fully to *all* publicly available information, so it should not matter whether the restatement is in the headline. However, limited attention theory (Hirshleifer and Teoh 2003) predicts that the speed and completeness of price reactions are both reduced when information is disclosed in a less noticeable format that some investors may overlook. This would occur because equilibrium stock prices reflect a weighted average of the beliefs of investors who attend to different signals, with weights that depend on the size and risk tolerance of each investor group (Hirshleifer et al. 2004). Thus, price reactions to restatements that are omitted from the press release headline (for example, restatements disclosed in an earnings press release) may be smaller and/or extend over a longer period than headline announcements. We therefore test the following null hypothesis:

> **H1:** The level of prominence given to an accounting restatement in the press release does not affect the magnitude of the market's response.

Disclosure prominence may also affect the likelihood of shareholder litigation. This could occur if potential litigants are themselves subject to limited attention. For example, class action law firms typically have staff assigned to clip newspaper articles about corporate events with lawsuit potential, and a headlined event is less likely to be overlooked. Further, class action lawyers can more easily link a large price decline to managers' actions when the news is prominently disclosed. We test the following null hypothesis:

> **H2:** The level of prominence given to an accounting restatement in the press release does not affect the likelihood of class action litigation.

Press release prominence could also have a second-order effect on litigation via returns if greater prominence leads to a more sizable price decline, which then exceeds a threshold that draws litigators' attention. We therefore control for announcement-period returns in testing H2. Finally, managers' assessment of the likelihood of a lawsuit could affect their selection of the level of press release prominence. Rogers and Van Buskirk (2008) provide evidence suggesting that managers believe reduced disclosure lowers the likelihood of another lawsuit. Specifically, they find that firms reduce the frequency of management forecasts (and conference calls) in the year following a lawsuit filing. Accordingly, we control for potential endogeneity in our tests.

III. SAMPLE SELECTION AND DATA DESCRIPTION

Our sample is taken from the GAO (2003) database, which lists 919 companies that announced an accounting restatement between January 1, 1997 and June 30, 2002. The GAO report based on this data (2002) states that the database excludes restatements arising from normal corporate activity or simple presentation issues.[6] The GAO tracks restatements based on the press release announcement date, not the date on which adjusted financial statements are filed. In addition to the company name and announcement date, the database reports the ticker symbol, the entity that prompted the restatement (e.g., auditor, company, SEC, unknown), and the accounting issue(s) selected from nine categories. The GAO list of restating companies is important because it influenced passage of the Sarbanes-Oxley Act of 2002 by showing Congress that accounting abuses were widespread during the market bubble of the 1990s. About 10 percent of all companies listed on the NYSE, NASDAQ, and Amex have a restatement cited in this database (GAO 2002, 16).

Table 1 reconciles our sample to the GAO list. We delete 423 of the 919 companies because they are not covered by CRSP and Compustat and another 77 firms that do not have short-interest data. Next, we exclude 30 firms for which we could not locate the press release.[7] Finally, to reduce the influence of outliers, we omit eight firms whose returns are in the top or bottom 1 percent of the cumulative market-adjusted returns distribution. The final sample consists of 381 restatement firms. In untabulated results, we find that restating firms are widely distributed across 49 two-digit industries with only modest differences from the Compustat population. Specifically, business services provides 15.49 percent of the restatement sample (59 of 381 companies), which is higher than the industry's 11.01 percent representation in Compustat. Only five other industries provide more than 12 sample firms.

TABLE 1
Sample Selection

	Number of Observations
Restatement firms in General Accounting Office study[a]	919
Less: Firms without CRSP or Compustat data	(423)
Less: Firms without short-selling information[b]	(77)
Less: Firms without press release information available	(30)
Less: Possible outliers[c]	(8)
Final Sample	381

[a] The General Accounting Office (GAO 2002) prepared a report for the U.S. Senate Committee on Banking, Housing, and Urban Affairs, chaired by Senator Sarbanes. The underlying database (GAO 2003) identifies 919 firm restatements from January 1, 1997 to June 30, 2002.

[b] Short-selling information is taken from a proprietary database that contains short interest for firms listed on the NYSE, AMEX, and Nasdaq.

[c] Observations whose cumulative abnormal returns around the restatement date fall in the top or bottom 1 percent of the distribution are deleted as possible outliers.

---

[6] The GAO (2002, 2) further states that their sample includes so-called "aggressive" accounting practices, intentional and unintentional misuse of facts applied to financial statements, oversight or misinterpretation of accounting rules, and fraud. An important advantage of the GAO list is that the researcher does not make the judgment about which restatements constitute aggressive accounting practices. A second advantage is that the list is publicly available, so other researchers can replicate and extend research results. An important disadvantage is that some of the restatements appear to be unintentional technical errors, rather than serious accounting irregularities.

[7] To ensure that we are capturing management's disclosure choices, we use press releases sent out over either PR Newswire or Business Wire, when available. These services distribute the exact press release issued by a firm, while other media outlets may alter the press release to conform to their format guidelines, space requirements, etc. Almost all of the press releases in our sample are from one of these two sources.

Table 2 provides a cross-tabulation of accounting issue with the three categories of disclosure prominence. Because some firms restate for more than one accounting issue, the 442 accounting issues exceed the 381 companies in the restatement sample. Revenue recognition is the most common type of accounting misstatement for each of the three prominence categories. Of particular note, a large portion of the companies providing footnote-only disclosure restate revenue (25 of 37). This is the primary reason that the 2.17 prominence score for revenue restatements is below the full-sample average of 2.33. Later in the paper, we report robustness tests to determine the effect of revenue restatements on our findings.

IV. MULTIVARIATE MODELS

We use OLS regression to investigate whether disclosure prominence is associated with returns around the restatement announcement (H1) and use logistic regression to examine whether disclosure prominence influences the likelihood of class action litigation (H2). For each hypothesis, we control for factors other than press release prominence that could influence the results.

TABLE 2

Cross-Tabulation of Accounting Issue by Press Release Prominence

Coded As	High Prominence (3)	Medium Prominence (2)	Low Prominence (1)	Frequency[a]	Average Prominence Scorce[b]
Revenue	52	81	25	158	2.17
Cost or Expense	28	42	3	73	2.34
Restructuring, Assets or Inventory	26	24	3	53	2.43
Other	17	13	2	32	2.47
Acquisitions and Mergers	14	13	1	28	2.46
Securities Related	12	12	0	24	2.50
IPR&D	4	15	2	21	2.10
Reclassification	8	10	1	19	2.37
Related-Party Transactions	6	7	0	13	2.46
Unspecified	6	3	0	9	2.67
Loan-Loss	6	1	0	7	2.86
Tax-Related	2	3	0	5	2.40
Total	181	224	37	442	2.33

This table describes the various accounting issues underlying each restatement in our sample. It includes the total frequency of each issue, along with the frequency across high, medium, and low prominence disclosure groups.

[a] Each restatement can be caused by one or more accounting issues. Therefore, the total number of accounting issues (442) is greater than the number of firms in our sample (381).

[b] Disclosure emphasis is coded as 3 (high prominence) for any press release that mentions the earnings restatement in the headline; coded 2 (medium prominence) for press releases that do not mention the restatement in the headline, but that discuss the restatement in some detail within the body of the report; and coded 1 (low prominence) for press releases that only mention the restatement in the footnotes of the report. The Average Prominence Score = [(# High Prominence) * (3) + (# Medium Prominence) * (2) + (# Low Prominence) * (1)]/# Total.

Abnormal Returns Surrounding Restatement Announcements

We estimate the following OLS regression model to test our first hypothesis:

$$\begin{aligned} CAR = {} & \alpha + \beta_1 PROMINENCE\ INDEX + \beta_2 RESTATEMENT\ MAGNITUDE \\ & + \beta_3 MAGNITUDE\ SOURCE + \beta_4 IRREGULARITY \\ & + \beta_5 SHORT\ INTEREST + \beta_6 SEC + \beta_7 REVENUE \\ & + \beta_8 RULE\ CHANGE + \beta_9 COUNT + \beta_{10} MKTCAP \\ & + \beta_{11} ANALYST + \beta_{12} INSTITUTIONAL\ OWN + \beta_{13} FINANCIAL + \varepsilon, \quad (1) \end{aligned}$$

where *CAR* is the cumulative abnormal return, calculated as the raw stock return minus the CRSP equally weighted market return.[8] *CAR* is measured over a three-day window centered on the restatement announcement date (−1, +1). In additional analyses, we extend the window to 20 days after the announcement date (−1, +20). As discussed in developing the hypotheses, we code *PROMINENCE INDEX* as 3 for headline disclosure of the restatement in the press release, 2 for discussions of the misstatement without a headline citing the restatement, and 1 for footnote-only disclosure.

We use three measures to control for differences in misstatement severity. The first measure, *RESTATEMENT MAGNITUDE*, is the cumulative earnings effect of the restatement scaled by total assets. Specifically, we hand-collect information on the cumulative overstatement or understatement of prior earnings as reported in company press releases (where available). We scale this dollar figure by total assets measured at the fiscal year-end prior to the restatement announcement and then winsorize at the 5th and 95th percentiles. We collect the restatement magnitude from SEC filings for 111 firms that do not report this information in the press release. Consistent with prior research (Palmrose et al. 2004; Lev et al. 2008), we then include an indicator variable, *MAGNITUDE SOURCE*, to allow for the possibility that magnitudes that are only disclosed after the press release announcement are systematically higher or lower.[9]

The second measure uses a classification scheme developed by Hennes et al. (2008) to distinguish between an "error" and an "irregularity." A dichotomous variable, *IRREGULARITY*, takes on the value of 1 if any of the following occurs: (1) any variation of the words "fraud" or "irregularity" is used in the press release; (2) the misstatement leads to an SEC Accounting and Auditing Enforcement Action (AAER); or (3) the company initiates an independent (non-SEC) investigation into the accounting misstatement.[10] Restatements in which none of the above occurs are classified as errors, and *IRREGULARITY* is coded as 0. The third severity measure, *SHORT INTEREST*, is based on the finding in Efendi et al. (2005) that the level of short interest increases substantially prior to the announcement of a severe accounting misstatement. This result suggests that short sellers have value-relevant private information about accounting quality that could overlap with the information managers use to select press release prominence. We therefore use short interest as a proxy for private information. *SHORT INTEREST* is the number of shares sold short expressed as a percentage of total shares outstanding, as calculated at the end of the month prior to the restatement announcement. Detailed definitions of all our variables are in the Appendix.

---

[8] In additional tests, we use Fama-French three-factor returns (Fama and French 1993) and value-weighted returns as the dependent variable in our model. Results are unchanged.

[9] As an alternative to using magnitude information collected from SEC filings, we also use conditional mean imputation (Allison 2002) to infer amounts that are not announced in the initial press release and substitute the imputed amounts for the missing data. The results are unchanged in both our returns and litigation models. Deleting the 111 observations also does not change our results.

[10] Following Hennes et al. (2008), "independent investigations" are those conducted by an (independent) audit committee, by a special committee of outside directors, by an outside forensic firm, or by an independent auditing firm that is not the usual auditor.

We also include four variables that have been used in prior studies to capture characteristics of the restatement (Lev et al. 2008; Palmrose and Scholz 2004; Palmrose et al. 2004; Wu 2003). *SEC* takes on the value of 1 if the press release mentions an SEC inquiry, and 0 otherwise. Based on prior research, we expect a positive sign on the *SEC* coefficient (i.e., a smaller price decline at the announcement). The rationale is that the market would usually know a restatement is forthcoming when it is important enough to attract SEC attention, and the actual announcement resolves uncertainty about the extent of a known accounting problem. *REVENUE* is coded 1 if the restatement is at least partly due to revenue recognition problems, and 0 otherwise. We expect a negative sign on the *REVENUE* coefficient because prior research finds that investors react more negatively to revenue restatements than to those caused by other accounting issues. *RULE CHANGE* takes on the value of 1 for firms that cite an accounting pronouncement as the reason for their restatement, and 0 otherwise. We expect the coefficient to be positive because these restatements are caused by either a change in rules or a new interpretation of existing accounting rules, rather than by a clear violation of GAAP. The fourth variable, *COUNT*, records the number of distinct accounting issues cited in the GAO database (2003). We expect the market to react more negatively to restatements stemming from multiple GAAP violations.

We also include some general control variables. *MKTCAP* is the natural log of the firm's market capitalization measured at the fiscal year-end prior to the restatement announcement. *ANALYST* is the number of analysts with a forecast included in the last consensus earnings forecast available prior to the restatement announcement. *INSTITUTIONAL OWN* is the level of institutional ownership at the end of the quarter prior to the restatement announcement, scaled by total shares outstanding on the same date. High levels of analyst following or institutional ownership are likely to expedite price reactions to the restatement announcement. Finally, we include a dummy variable for firms operating in the financial industry (SIC 6000–6999). Based on prior research, we anticipate less negative price reactions for *FINANCIAL* firms.

Endogeneity is a potential concern in modeling the relation between market returns and press release prominence. We include measures of misstatement severity to control for endogeneity that would arise if managers select disclosure prominence based on severity; and we use the Heckman (1979) two-stage procedure to address other sources of endogeneity (such as the firm's openness in communicating with shareholders). In the first stage, we obtain three inverse Mills ratios from an ordered Probit model predicting *PROMINENCE INDEX*, and in the second stage, we include the inverse Mills ratios in the returns model. The Heckman procedure requires that we include all the variables used to explain returns, along with at least one additional variable that could affect the level of prominence managers choose but that is not likely to affect announcement date returns. We use the number of management earnings forecasts in the year prior to the restatement announcement as the additional variable in the prominence choice model. We report results both with and without this endogeneity adjustment.

Disclosure Prominence and the Likelihood of Class Action Lawsuits

We investigate the association between disclosure prominence and the likelihood of shareholder litigation by estimating the following logistic regression model:

$$\begin{aligned} LITIGATION = {} & \alpha + \beta_1 PROMINENCE\ INDEX \\ & + \beta_2 RESTATEMENT\ MAGNITUDE \\ & + \beta_3 MAGNITUDE\ SOURCE + \beta_4 IRREGULARITY \\ & + \beta_5 SHORT\ INTEREST + \beta_6 CONCURRENT\ RETURN \\ & + \beta_7 SEC + \beta_8 REVENUE + \beta_9 RULE\ CHANGE \\ & + \beta_{10} COUNT + \beta_{11} PRIOR\ STOCK\ RETURN \\ & + \beta_{12} POST\ STOCK\ RETURN + \beta_{13} SHARETURNOVER \\ & + \beta_{14} MKTCAP + \beta_{15} ANALYST + \beta_{16} INSTITUTIONAL\ OWN \\ & + \beta_{17} FINANCIAL + \varepsilon, \end{aligned} \tag{2}$$

where *LITIGATION* is equal to 1 if a class action lawsuit is filed against a firm in response to their restatement, and 0 otherwise. We predict that more prominent disclosure will increase the likelihood of litigation. Thirteen of the 17 variables in model 2 are also used in model 1 to predict returns. The hypothesized coefficient signs for the litigation model are generally opposite to those predicted in the returns model because we expect variables that are negatively (positively) associated with returns to lead to an increase (decrease) in litigation.

We include four new variables in the litigation model. Prior research shows that *SHARE TURNOVER* is positively associated with the likelihood of litigation. The explanation is that when shares trade rapidly, more investors are potentially harmed by any mispricing, thus increasing the number of potential plaintiffs (Field et al. 2005; Gande and Lewis 2009). We calculate this variable using the formula in Field et al. (2005): $[1 - \Pi_t(1 - \mathit{volume\ traded}_t / \mathit{total\ shares}_t)]$, accumulated from daily trading volume (for each day *t*) over the one-year period ending two days prior to the restatement announcement date. The three other new variables measure changes in stock price. *CONCURRENT RETURN* is the compounded raw return over the three-day period centered on the date of the restatement announcement (−1, +1). We expect the likelihood of litigation to increase as the share price declines (Palmrose and Scholz 2004). *PRIOR STOCK RETURN* is the compounded raw return over the one-year period ending two days before the restatement announcement (−252, −2). Weak stock price performance in the recent past has been shown to increase the incidence of lawsuits (Jones and Weingram 1996; Gande and Lewis 2009). *POST RETURN* is the compounded raw return over the (+2, +20) interval following the restatement announcement. This return window allows for a delayed market reaction as investors assess the seriousness of the accounting misstatement.[11]

Endogeneity concerns also arise in our litigation model because managers may choose a level of prominence to reduce litigation risk. In addition to the misstatement severity measures, we again use the Heckman (1979) procedure to control for this possibility. In the first stage, we use an ordered Probit regression to predict *PROMINENCE INDEX*. The model includes each variable used to explain litigation, along with the number of management earnings forecasts in the prior year as our instrumental variable. We report coefficient values both with and without the effect of the inverse Mills ratios.

---

[11] We also re-run our models using abnormal returns rather than raw returns and our results are virtually identical. We choose to report raw returns to remain consistent with prior research (Field et al. 2005; Ferris and Pritchard 2001).

V. RESULTS FOR MARKET RETURNS

Returns over Short and Longer Windows

Table 3, Panel A, reports mean abnormal returns for a three-day window centered on the announcement date (−1, +1), for a longer window (−1, +20), and for the difference between those two windows (drift). In the three-day window, abnormal returns for high, medium, and low prominence are −8.3 percent, −4.0 percent, and −1.5 percent, respectively. The returns for high and medium prominence are significantly different from zero. Paired comparisons show that the return difference between high and either medium or low prominence disclosers are statistically significant at the 0.01 level, with the difference between medium and low categories significant at the 0.10 level.

In the longer window (−1, +20), the overall return declines to −6.7 percent (from −5.5 percent), with average returns of −7.9 percent, −6.4 percent, and −3.2 percent for the high, medium, and low prominence disclosure categories. Each of those returns is significantly different from zero. Paired comparisons show that the differences between high and low prominence firms and between medium and low prominence firms are still statistically significant. Importantly, the return difference between high and medium prominence firms is now a statistically insignificant −1.5 percent. If we assume that returns fully reflect the market reaction to restatements in the longer window, this result suggests that investors respond similarly to high and medium prominence restatements. This observation is supported by the short-to-long-window return drift of −2.4 percent (significant at the 0.10 level) for firms providing medium disclosure and +0.4 percent for firms providing high prominence disclosure. The net drift of 2.8 percent, which is significant at the 0.10 level, substantially reduces the return difference between high and medium prominence firms observed in the short-window. This result provides preliminary evidence for rejecting H1 (i.e., prominence does not affect the market's response).

Understanding the Post-Announcement Drift

The drift results reported in Table 3, Panel A, suggest that market participants initially underestimate the importance of restatements disclosed with medium prominence and, to a lesser extent, low prominence, but subsequently correct their under-reaction. This pattern could occur because investors respond to new information disclosed after the restatement. Alternatively, investors may further evaluate the original press release information, which would suggest that the initial under-reaction is due to limited attention (Hirshleifer and Teoh 2003). To evaluate these possibilities, we examine corporate press releases and other news items disclosed in the month after the announcement.

As reported in Table 3, Panel B, we find news items for 81 of the 224 firms that provide either medium or low prominence disclosure, and no news items for the remaining 143 firms. We classify the news items into three categories: (1) news about a filed or pending lawsuit related to the restatement, (2) other news related to the restatement, or (3) news unrelated to the restatement. For the lawsuit and unrelated news categories, a statistically insignificant negative drift of about 2 percent occurs. For the category of other news related to the restatement, the drift is significant (at the 0.10 level) and larger at 7.7 percent. Further

TABLE 3
Mean Abnormal Returns around the Restatement Announcement

Panel A: All Restating Firms

Prominence Index	Number of Firms	Short Window (−1, +1)[a]	Longer Window (−1, +20)[a]	Drift (see Panel B)[a]
High (3)	157	−8.3%***	−7.9%***	+0.4%
Medium (2)	189	−4.0%***	−6.4%***	−2.4%*
Low (1)	35	−1.5%	−3.2%*	−1.7%
Full Sample	381	−5.5%***	−6.7%***	−1.2%
Paired Comparisons				
High (3) versus Low (1)		−6.8%***	−4.7%**	+2.1%
High (3) versus Medium (2)		−4.3%***	−1.5%	+2.8%*
Medium (2) versus Low (1)		−2.5%*	−3.2%*	−0.7%

Panel B: Analysis of Drift for Medium and Low Prominence Firms

	Drift during Month after Restatement		
News Items during Month after Restatement	Medium Prominence[a]	Low Prominence[a]	Total[a]
Lawsuit filed or pending	−2.6% (n = 29)	NA (n = 0)	−2.6% (n = 29)
Other news related to restatement	7.7%* (n = 22)	NA (n = 0)	7.7%* (n = 22)
Other news unrelated to restatement	−2.6% (n = 21)	0.3% (n = 9)	−1.7% (n = 30)
All firms with news	0.6% (n = 72)	0.3% (n = 9)	0.5% (n = 81)
Firms with no news (see Panel C)	−4.0%*** (n = 117)	−2.3% (n = 26)	−3.7%*** (n = 143)

Panel C: Influence of Analyst Following on Drift for Firms with No News

	Drift during Month after Restatement		
Analyst Following	Medium Prominence[a]	Low Prominence[a]	Total[a]
Zero analysts	−0.9% (n = 12)	NA (n = 0)	−0.9% (n = 12)
One or two analysts	−1.7% (n = 45)	−1.6% (n = 8)	−1.7% (n = 53)
Three or more analysts	−6.4%*** (n = 60)	−2.5% (n = 18)	−5.5%*** (n = 78)
All firms with no news items	−4.0%*** (n = 117)	−2.3%* (n = 26)	−3.7%*** (n = 143)

---

*, **, *** Indicate a significant difference in a two-tailed t-test with p-values of 0.10, 0.05, and 0.01, respectively.

This table reports mean abnormal returns by prominence level in various windows around the restatement announcement date and in the subsequent month.

[a] Short window and Longer window returns are cumulative abnormal returns (CARs), calculated as the raw stock return minus the CRSP equally weighted market portfolio return, with day 0 indicating the restatement announcement.

investigation reveals that 14 of the 22 firms experienced a positive return, and ten of these 14 firms reported additional information about the restatement to the SEC during the month (on form 8-K, 10-K/A, or 10-Q/A). The positive return for the additional news category appears to occur because the information reduces uncertainty about the scope of the accounting problem.

By comparison, the 143 firms with no additional news in the month following the restatement display a negative price drift of −3.7 percent (statistically significant at the 0.01 level). This negative post-announcement drift is consistent with a delayed market reaction to the disclosures in the earlier press release. To provide additional insight into this result, we examine the role of analyst following. As reported in Table 3, Panel C, we find a significant and sizable negative drift of −5.5 percent for firms that provide medium or low prominence disclosure and that are followed by at least three analysts. In contrast, there is a small, insignificant negative drift for firms with no analyst following (−0.9 percent) or that are followed by one or two analysts (−1.7 percent).[12]

To summarize, we find evidence of a post-announcement negative drift in prices for firms providing less prominent press release disclosure of their restatement. Much of the post-announcement drift occurs for firms without additional news, which is consistent with limited attention by investors. Higher analyst following seems to be the catalyst for much of the post-announcement price correction by these firms.

Multivariate Analysis

Pairwise Correlations between CAR and Independent Variables

Short-window returns around restatements (*CAR*) are significantly negatively associated with the level of press release prominence (*PROMINENCE INDEX*), with the restatement amount (*RESTATEMENT MAGNITUDE*), and with restatements classified as an irregularity rather than an error (*IRREGULARITY*). Returns are also significantly lower for revenue restatements (*REVENUE*) but are higher for restatements arising from an accounting rule change (*RULE CHANGE*). Each of these five associations is consistent with our expectations (results untabulated).

Abnormal Returns Surrounding Restatement Announcements

Table 4 presents regression results for short-window cumulative abnormal returns. We report one version of the model without an adjustment for endogeneity and a second version with inverse Mills ratios. We begin our discussion with the model excluding inverse Mills ratios. The most important finding is that the *PROMINENCE INDEX* is significant at the 0.003 level with a coefficient of −0.034, which is similar in magnitude to the univariate results reported in Table 3. Announcement day returns are thereby reduced by about 3.4 percent for each change in prominence level (from high to medium, or from medium to low prominence).

The returns model includes variables to control for misstatement severity. The first measure, *RESTATEMENT MAGNITUDE*, is significant at the 0.034 level and the coefficient estimate of 0.456 indicates that a negative restatement of 10 percent of total assets (which would correct overstated earnings) would result in incremental negative returns of about 4.6 percent. A related variable, *MAGNITUDE SOURCE*, which takes on the value of 1 if

---

[12] One explanation for this result is that analysts scrutinize press releases in the weeks after the restatement announcement and communicate their findings to investors by updating their forecasts or through other means. The likelihood that analysts uncover new information increases with analyst following, as does the number of investors to whom the new information is communicated. Both of these influences may contribute to the more sizable negative post-announcement drift we observe for firms covered by at least three analysts.

TABLE 4
Multivariate Analysis of Abnormal Returns around the Restatement Announcement

Short Window *CAR* (−1, +1)

$$
\begin{aligned}
CAR = \alpha &+ \beta_1 PROMINENCE\ INDEX + \beta_2 RESTATEMENT\ MAGNITUDE \\
&+ \beta_3 MAGNITUDE\ SOURCE + \beta_4 IRREGULARITY + \beta_5 SHORT\ INTEREST \\
&+ \beta_6 SEC + \beta_7 REVENUE + \beta_8 RULE\ CHANGE + \beta_9 COUNT + \beta_{10} MKTCAP \\
&+ \beta_{11} ANALYST + \beta_{12} INSTITUTIONAL\ OWN + \beta_{13} FINANCIAL + \varepsilon
\end{aligned}
$$

		Full Model			Full Model with Inverse Mills Ratios[a]		
Variable	**Prediction**	**Coefficient Estimate**	**t-stat**	**p-value[b]**	**Coefficient Estimate**	**t-stat**	**p-value[b]**
Intercept	?	0.059	1.323	0.187	0.200	0.267	0.790
PROMINENCE INDEX	?	−0.034	−3.028	0.003	−0.038	−3.305	< 0.001
Misstatement Severity							
RESTATEMENT MAGNITUDE	(+)	0.456	1.837	0.034	0.921	2.768	0.003
MAGNITUDE SOURCE	?	0.008	0.553	0.580	0.049	2.060	0.040
IRREGULARITY	(−)	−0.034	−1.844	0.033	−0.060	−2.674	0.004
SHORT INTEREST	(−)	−0.001	−0.455	0.325	0.001	0.508	0.306
Restatement Characteristics							
SEC	(+)	0.034	1.645	0.050	0.015	0.650	0.258
REVENUE	(−)	−0.056	−3.926	< 0.001	−0.043	−2.712	0.004
RULE CHANGE	(+)	0.041	2.295	0.011	0.103	2.961	0.002
COUNT	(−)	0.006	0.362	0.359	0.017	1.012	0.156
Other Controls							
MKTCAP	?	−0.008	−1.516	0.130	0.004	0.582	0.561
ANALYST	?	0.000	0.271	0.787	−0.001	−0.612	0.541
INSTITUTIONAL OWN	?	0.040	1.297	0.195	0.016	0.500	0.617
FINANCIAL	(+)	0.032	1.722	0.043	−0.004	−0.167	0.434
n		353[c]			353[c]		
Adjusted R^2		10.20%			11.22%		

---

[a] To control for potential endogeneity between managers' disclosure choice (proxied for by *PROMINENCE INDEX*) and returns at the disclosure date, we run a first-stage ordered probit model predicting *PROMINENCE INDEX* and generate three inverse Mills ratios. These ratios are included in the regression above. As their coefficients are not significant under any specification, and for the sake of brevity, we omit their coefficients from the table.

[b] Two-tailed tests (one-tailed tests) are shown for variables without (with) a signed prediction.

[c] 23 firms were excluded from the regression due to missing market capitalization information and five were excluded due to missing restatement magnitude information.

Variable Definitions:

CAR = cumulative abnormal return, calculated as the raw stock return minus the CRSP equally weighted market portfolio return measured over the (−1,+1) window, with day 0 indicating the restatement announcement;

PROMINENCE INDEX = 3 (high prominence) for any press release that mentions the earnings restatement in the headline; 2 (medium prominence) for press releases that do not mention the restatement in the headline, but that discuss the restatement in some detail within the body of the report; 1 (low prominence) for press releases that only mention the restatement in the footnotes of the report;

(*continued on next page*)

TABLE 4 (continued)

RESTATEMENT MAGNITUDE = cumulative earnings effect of the restatement scaled by total assets measured as of the fiscal year-end prior to the restatement announcement, winsorized at the 5th and 95th percentiles. When available, this data was collected from the initial press release announcing the restatement. When not available, it was collected from subsequent SEC filings;
MAGNITUDE SOURCE = 1 if the restatement magnitude was not included in the press release initially announcing the restatement (i.e. the information was missing), and 0 otherwise;
IRREGULARITY = 1 if any of the following applies to the restatement: (a) variants of the words "fraud" or "irregularity" are used in the restatement announcement, (b) the firm is subject to an SEC Auditing and Accounting Enforcement Action, or (c) the firm initiates an independent investigation into the accounting issue, and 0 otherwise;
SHORT INTEREST = percent of short interest [100 * (number of shares sold short)/(total shares outstanding)] as of the month prior to the restatement announcement;
SEC = 1 if the press release indicates an SEC inquiry, and 0 otherwise;
REVENUE = 1 if any part of the restatement is due to revenue recognition problems, and 0 otherwise;
RULE CHANGE = 1 if a change in accounting rules was mentioned in the press release as the reason for the restatement; and 0 otherwise;
COUNT = number of distinct accounting issues;
MKTCAP = natural log of market capitalization (closing stock price * common shares outstanding) of the firm measured as of the end of the fiscal year prior to the restatement;
ANALYST = number of analysts whose earnings forecasts are included in the consensus formed prior to the most recent quarterly or annual earnings announcement prior to the restatement announcement;
INSTITUTIONAL OWN = level of institutional ownership at the end of the quarter prior to the restatement announcement, scaled by total shares outstanding on the same date; and
FINANCIAL = 1 if the firm operates in the financial services sector (SIC codes 6000–6999), and 0 otherwise.

we obtain the restatement amount from an SEC filing (rather than the press release), is not significant. The second severity measure, *IRREGULARITY*, is significant at the 0.033 level and the coefficient indicates returns would be lower by 3.4 percent for an irregularity than for an error. The third severity measure, *SHORT INTEREST*, does not have a significant effect on announcement returns.

Three restatement characteristics are also significantly associated with market returns. Restatements that have been previously discussed with the SEC have a positive (less negative) return of 3.4 percent (significant at the 0.05 level). The explanation (from prior research) is that investors knew an accounting problem existed and information in the press release reduces uncertainty about the nature of the problem. The *REVENUE* coefficient is significant at less than the 0.001 level, and its value indicates that returns are lower by an average of 5.6 percent when revenue is not recognized in accordance with GAAP. The *RULE CHANGE* coefficient is significant at the 0.011 level. Restatements that arise from a change in rules or a reinterpretation of existing accounting rules, rather than from a clear violation of GAAP, have a positive (less negative) return of 4.1 percent. Of the general control variables, only *FINANCIAL* is significantly associated with returns. The positive and significant coefficient indicates that returns are more positive (less negative) by 3.2 percent for restatements reported by banks and other financial firms.

The three columns on the right side of Table 4 show the results for our full model when we include the inverse Mills ratios to control for potential endogeneity.[13] The most important feature of this model is that, after controlling for endogeneity, *PROMINENCE INDEX* is still significantly associated with abnormal returns. The coefficient increases by about 10 percent from −0.034 to −0.038. The coefficients of the misstatement severity measures, *RESTATEMENT MAGNITUDE* and *IRREGULARITY,* are again highly significant and are considerably larger than in the versions of the model that do not control for endogeneity. The *MAGNITUDE SOURCE* coefficient is also significant, indicating that firms not reporting an estimate of the magnitude in the press release have significantly higher (less negative) returns of 4.9 percent on average. Examining the remaining variables, the *REVENUE* coefficient is smaller but still significant; the *RULE CHANGE* coefficient is much larger and still significant; and the *SEC* and *FINANCIAL* coefficients become insignificant.

To further explore the negative return drift for medium prominence firms reported in Table 3, we also run a model explaining longer-window (−1, +20) returns and find that the coefficient on *PROMINENCE INDEX* is no longer significant (untabulated). This analysis verifies that although less prominent disclosures are rewarded with less negative returns in the short-run, investors correct their under-reaction in the long-run. Restatement severity (*RESTATEMENT MAGNITUDE* and *IRREGULARITY*) is still associated with returns in the long-window.

In sum, the *PROMINENCE INDEX* coefficient is negative and highly significant in both versions of the model explaining short (−1, +1) window returns. Companies that disclose a restatement without mentioning it in the headline of the press release thus experience less negative announcement returns. Based on this finding and the negative price drift (discussed previously), we reject null H1, which states that the prominence given to an accounting restatement in the press release does not affect the magnitude of the market's response. As discussed in the next section, this conclusion does not change after robustness tests.

Robustness Tests

The announcement of an accounting restatement often coincides with the release of current earnings news. To ensure that our results are not due to return responses to the earnings surprise, we re-estimate our model on a reduced sample of firms that announce earnings in the same three-day window as their restatement announcement. This occurs for 173 firms in our sample (45 percent of the sample), comprised of 11 percent of the high prominence firms, 40 percent of the medium prominence firms, and 74 percent of the low prominence firms.

We calculate earnings surprise (*SURPRISE*) as the difference between actual quarterly earnings per share and the most recent consensus analyst earnings forecast issued prior to the earnings announcement (both reported by I/B/E/S), scaled by stock price measured just before the forecast date. *BEAT* is an indicator variable for firms whose earnings exceed analyst expectations, defined as 1 if *SURPRISE* > 0, and 0 otherwise. Requiring consensus analyst forecasts and stock prices reduces our subsample to 120 firms.

---

[13] In the first stage model (untabulated) that predicts *PROMINENCE INDEX*, we find the number of management forecasts is significantly associated with *PROMINENCE INDEX* (negative sign). Other significant variables are *MAGNITUDE* (negative sign), *MAGNITUDE SOURCE* (negative sign), *IRREGULARITY* (positive sign), *RULE CHANGE* (negative sign), *MKTCAP* (negative sign), and *FINANCIAL* (positive sign). In the second stage model, we find that the individual coefficients for the three inverse Mills ratios are statistically insignificant, so we do not report them on Table 5. A joint test of the three coefficients reveals that they are only marginally significant as a group.

Table 5 presents the results with *SURPRISE* and *BEAT* in our full model, both without and with the effects of the inverse Mills ratios. In both versions of the model, *SURPRISE* is positively and significantly associated with abnormal returns. The coefficient on *BEAT* is also significant, indicating that firms with earnings that exceed analyst expectations experience a substantially higher return than those that do not. More importantly, the *PROMINENCE INDEX* coefficient remains negative and highly significant, even though we estimate the model using less than one-third of our original sample.

As a further robustness check, we also add the earnings surprise controls to a model estimated on our full sample of restating firms, as opposed to the subsample discussed above (not tabulated). Firms that do not announce earnings are assigned a value of 0 for both *SURPRISE* and *BEAT*. Our findings do not change—the *PROMINENCE INDEX* coefficient remains highly significant. The coefficients on the other variables are of the same sign and significance as in Table 4, except for the *SEC* and *SURPRISE* coefficients, which are no longer significant.

Table 2 shows that the average prominence score for revenue restatements is considerably below the sample average (2.17 versus 2.33). Since revenue restatements are the most common type of accounting issue (158 of 442 instances), we investigate whether our results change if we run our model on a sample (1) excluding revenue restatement observations, or (2) including only revenue restatement observations. In untabulated results, we find that our main variable of interest, *PROMINENCE INDEX*, remains highly significantly associated with returns under both specifications.

Some prior studies view restatements initiated by an entity other than the company as more severe (Palmrose et al. 2004; Efendi et al. 2007). We therefore examine the effect of adding a variable for the entity that prompts the restatement, using the initiator field in the GAO database (2003), but we find that the coefficient on this indicator variable is insignificant.

Finally, some firms have more than one restatement included in the GAO database. As a sensitivity test, we exclude those observations and our results are not changed.

VI. LITIGATION RESULTS

We obtain information about litigation from the Securities Class Action Clearinghouse (http://securities.stanford.edu/). This website covers more than 2,500 issuers named in a federal class action securities fraud lawsuit since passage of the Private Securities Litigation Reform Act of 1995. It includes information about complaints, briefs, filings, and other litigation-related materials. For each sample firm, we reviewed this information to identify any litigation arising from the accounting restatement.

Table 6 reports the frequency of class action litigation arising from accounting restatements. Of the 381 companies in our sample, 73 (19 percent) are served with a class action lawsuit. The litigation rate declines monotonically with disclosure prominence. The rate is 27 percent (42 of 157) for firms providing high prominence disclosure, 16 percent (31 of 189) for firms providing medium prominence disclosure, and 0 percent for the 35 firms providing low prominence disclosure. A Chi-square test of association shows that lawsuit incidence is significantly associated with disclosure prominence (at the 0.001 level). While the litigation rate is lower for medium than for high prominence, the 16 percent rate for medium prominence suggests that some managers use medium prominence disclosure for an accounting misstatement that plaintiff attorneys view as serious.

Pairwise correlations (untabulated) show that eight variables have a significantly positive association with *LITIGATION*: *PROMINENCE INDEX*, *SEC*, *SHORT INTEREST*, *SHARE TURNOVER*, *MKTCAP*, *IRREGULARITY*, *ANALYST*, and *INSTITUTIONAL OWN*.

TABLE 5
Multivariate Analysis of Abnormal Returns around the Restatement Announcement with Controls for the Earnings Surprise

Short Window *CAR* (−1, +1)

$$\begin{aligned} \mathit{CAR} = {} & \alpha + \beta_1 \mathit{PROMINENCE\ INDEX} + \beta_2 \mathit{RESTATEMENT\ MAGNITUDE} \\ & + \beta_3 \mathit{MAGNITUDE\ SOURCE} + \beta_4 \mathit{IRREGULARITY} + \beta_5 \mathit{SHORT\ INTEREST} \\ & + \beta_6 \mathit{SEC} + \beta_7 \mathit{REVENUE} + \beta_8 \mathit{RULE\ CHANGE} + \beta_9 \mathit{COUNT} + \beta_{10} \mathit{MKTCAP} \\ & + \beta_{11} \mathit{ANALYST} + \beta_{12} \mathit{INSTITUTIONAL\ OWN} + \beta_{13} \mathit{FINANCIAL} + \beta_{14} \mathit{SURPRISE} \\ & + \beta_{15} \mathit{BEAT} + \varepsilon \end{aligned}$$

		Model with Earnings Surprise			Model with Earnings Surprise and Inverse Mills Ratios[a]		
Variable	**Prediction**	**Coefficient Estimate**	**t-stat**	**p-value[b]**	**Coefficient Estimate**	**t-stat**	**p-value[b]**
Intercept	?	−0.028	−0.291	0.771	−0.357	−0.882	0.380
PROMINENCE INDEX	?	−0.051	−2.454	0.016	−0.051	−2.274	0.025
Misstatement Severity							
RESTATEMENT MAGNITUDE	(+)	−0.063	−0.116	0.454	0.611	0.635	0.264
MAGNITUDE SOURCE	?	0.005	0.210	0.834	0.038	0.796	0.428
IRREGULARITY	(−)	−0.091	−1.709	0.045	−0.077	−1.333	0.093
SHORT INTEREST	(−)	0.000	−0.126	0.450	0.001	0.178	0.430
Restatement Characteristics							
SEC	(+)	0.059	1.442	0.076	0.018	0.290	0.386
REVENUE	(−)	−0.068	−2.743	0.004	−0.060	−2.277	0.013
RULE CHANGE	(+)	0.039	1.478	0.071	0.058	1.632	0.053
COUNT	(−)	0.036	0.930	0.177	0.045	1.104	0.137
Other Controls							
MKTCAP	?	0.008	0.731	0.467	0.014	1.095	0.277
ANALYST	?	−0.001	−0.424	0.672	−0.002	−0.792	0.430
INSTITUTIONAL OWN	?	−0.030	−0.580	0.563	−0.051	−0.845	0.400
FINANCIAL	(+)	0.097	2.312	0.011	0.063	1.092	0.139
SURPRISE	(+)	4.390	2.153	0.017	3.962	1.826	0.036
BEAT	(+)	0.060	2.121	0.018	0.071	2.206	0.015
n		111[c]			111[c]		
Adjusted R^2		26.85%			25.14%		

---

[a] To control for potential endogeneity between managers' disclosure choice (proxied for by *PROMINENCE INDEX*) and returns at the disclosure date, we run a first-stage ordered probit model predicting *PROMINENCE INDEX* and generate three inverse Mills ratios. These ratios are included in the regression above. As their coefficients are not significant under any specification, and for the sake of brevity, we omit their coefficients from the table.

[b] Two-tailed tests (one-tailed tests) are shown for variables without (with) a signed prediction.

[c] This sample is restricted to firms that announce earnings and a restatement in the same three-day window and that have analyst forecast and stock price information. Seven firms were excluded from the regression due to missing market capitalization information and two with missing restatement magnitude information.

Variable Definitions:

SURPRISE = difference between actual quarterly earnings per share and the most recent consensus analyst earnings forecast issued prior to the earnings announcement, both reported by I/B/E/S, scaled by stock price measured just prior to the date of the forecast; and

BEAT = 1 if the firm beat analyst expectations (i.e. *SURPRISE* > 0), and 0 otherwise.

All other variables are defined in Table 4.

TABLE 6
Distribution of Class Action Lawsuits by Disclosure Prominence

	High Prominence	Medium Prominence	Low Prominence	Total
Class Action Lawsuit Filed	42	31	0	73
No Class Action Lawsuit Filed	115	158	35	308
Total	157	189	35	381
Percentage of Observations with Litigation	27%	16%	0%	19%
Chi-square[a]	15.065 (p = 0.001)			

This table presents the distribution of class action lawsuits across prominence category. Litigation information is from the Securities Class Action Clearinghouse website.

[a] The Chi-square test examines the association between lawsuit incidence and disclosure prominence. The significant Chi-square value indicates a difference in lawsuit initiation rates among prominence levels.

Two variables have a significant negative association: *CAR* and *RULE CHANGE*. The signs for each of these ten variables are in the hypothesized direction.

Table 7, Panel A, reports results for a logistic regression model with four types of misstatement severity measures. We find that the *PROMINENCE INDEX* coefficient is positive and significant at the 0.015 level. Taking the anti-log of the 0.746 coefficient estimate indicates that a reduction of one category in the *PROMINENCE INDEX* cuts the odds of a lawsuit roughly in half.[14]

Our study focuses on disclosure prominence, but we believe it is the first to provide evidence on factors that influence litigation for the GAO (2003) sample of restatements. Comparing the standardized coefficients, *SHARE TURNOVER* is the most influential factor, supporting the notion that litigators bring suits against heavily traded stocks because more investors can be included in the class action pool. Based on standardized coefficients, the next two variables of importance are measures of misstatement severity. The likelihood of litigation is higher for firms reporting an *IRREGULARITY* (rather than an error). Consistent with litigants using stock price declines to select litigation targets, *CONCURRENT RETURN* $(-1, +1)$ is also negatively associated with litigation. Each of these variables is significant at the 0.001 level. Restatements having a more negative effect on earnings (*MAGNITUDE*) are also more likely to result in litigation. In addition, litigation is more likely for firms subject to an SEC inquiry, for firms with a more negative *PRIOR STOCK RETURN* $(-252, -2)$, for *FINANCIAL* firms, and for those with higher levels of *INSTITUTIONAL OWN*. The latter result is likely a byproduct of the enactment of the Private Securities Litigation Reform Act of 1995, in which institutions were encouraged to take on the role of lead plaintiff in class action suits because they can best monitor the firm and the litigation proceedings (Cox and Thomas 2006). In a study by PricewaterhouseCoopers LLP (2003), 51 percent of all lead plaintiff positions in 2002 were occupied by institutional investors.

---

[14] Odds ratios are useful for the interpretation of individual coefficients but cannot be compared across variables as their value depends on the unit of measurement. For this reason, we do not report odds ratios in Table 7; instead we report standardized coefficients that measure changes in standard deviation units, therefore enabling us to compare the relative importance of each explanatory variable. Standardized coefficients, however, induce linearity in the model and should also be interpreted with care.

TABLE 7
Logistic Model of the Likelihood of Class Action Lawsuits following an Accounting Restatement

Panel A: Full Model

$$\begin{aligned} LITIGATION = {} & \alpha + \beta_1 PROMINENCE\ INDEX + \beta_2 RESTATEMENT\ MAGNITUDE \\ & + \beta_3 MAGNITUDE\ SOURCE + \beta_4 IRREGULARITY \\ & + \beta_5 SHORT\ INTEREST + \beta_6 CONCURRENT\ RETURN \\ & + \beta_7 SEC + \beta_8 REVENUE + \beta_9 RULE\ CHANGE + \beta_{10} COUNT \\ & + \beta_{11} PRIOR\ STOCK\ RETURN + \beta_{12} POST\ STOCK\ RETURN \\ & + \beta_{13} SHARE\ TURNOVER + \beta_{14} MKTCAP + \beta_{15} ANALYST \\ & + \beta_{16} INSTITUTIONAL\ OWN + \beta_{17} FINANCIAL + \varepsilon \end{aligned}$$

Variable	Prediction	Coefficient Estimates	Chi-Square p-value[a]	Standardized Coefficient
Intercept	?	−8.548	< 0.001	
PROMINENCE INDEX	(+)	0.746	0.015	0.264
Misstatement Severity				
RESTATEMENT MAGNITUDE	(−)	−13.777	0.021	−0.212
MAGNITUDE SOURCE	?	−0.002	0.996	−0.001
IRREGULARITY	(+)	2.120	< 0.001	0.429
SHORT INTEREST	(+)	0.039	0.183	0.103
CONCURRENT RETURN (−1, +1)	(−)	−5.377	< 0.001	−0.389
Restatement Characteristics				
SEC	(+)	0.830	0.051	0.144
REVENUE	(+)	−0.058	0.446	−0.016
RULE CHANGE	(−)	−0.729	0.129	−0.171
COUNT	(+)	0.243	0.282	0.054
Other Controls				
PRIOR STOCK RETURN (−252, −2)	(−)	−0.545	0.032	−0.194
POST STOCK RETURN (+2, +20)	(−)	0.757	0.200	0.080
SHARE TURNOVER	(+)	3.000	0.002	0.461
MKTCAP	?	0.061	0.679	0.066
ANALYST	?	−0.023	0.582	−0.074
INSTITUTIONAL OWN	?	2.500	0.009	0.356
FINANCIAL	?	1.118	0.065	0.223
n		343[b]		
−2 log likelihood		338.79		
Model Chi-square		138.94		
p-value		< 0.001		
Pseudo R^2		33%		
Correctly Classified [baseline of no litigation is 81%]		90.6%		

(*continued on next page*)

TABLE 7 (continued)

Panel B: Full Model with Inverse Mills Ratios[c]

$$\begin{aligned} LITIGATION = {} & \alpha + \beta_1 PROMINENCE\ INDEX + \beta_2 RESTATMENT\ MAGNITUDE \\ & + \beta_3 MAGNITUDE\ SOURCE + \beta_4 IRREGULARITY \\ & + \beta_5 SHORT\ INTEREST + \beta_6 CONCURRENT\ RETURN + \beta_7 SEC \\ & + \beta_8 REVENUE + \beta_9 RULE\ CHANGE + \beta_{10} COUNT \\ & + \beta_{11} PRIOR\ STOCK\ RETURN + \beta_{12} POST\ STOCK\ RETURN \\ & + \beta_{13} SHARE\ TURNOVER + \beta_{14} MKTCAP + \beta_{15} ANALYST \\ & + \beta_{16} INSTITUTIONAL\ OWN + \beta_{17} FINANCIAL \\ & + \beta_{18-20} INVERSE\ MILLS\ RATIOS + \varepsilon \end{aligned}$$

Variable	Prediction	Coefficient Estimates	Chi-Square p-value[a]	Standardized Coefficient
Intercept	?	23.007	0.533	
PROMINENCE INDEX	(+)	0.683	0.026	0.242
Misstatement Severity				
RESTATEMENT MAGNITUDE	(−)	−9.169	0.152	−0.141
MAGNITUDE SOURCE	?	0.351	0.607	0.089
IRREGULARITY	(+)	1.789	< 0.001	0.363
SHORT INTEREST	(+)	0.047	0.162	0.124
CONCURRENT RETURN (−1, +1)	(−)	−4.129	0.028	−0.299
Restatement Characteristics				
SEC	(+)	0.496	0.204	0.086
REVENUE	(+)	0.143	0.385	0.039
RULE CHANGE	(−)	−0.297	0.385	−0.070
COUNT	(+)	0.359	0.215	0.080
Other Controls				
PRIOR STOCK RETURN (−252, −2)	(−)	−0.612	0.023	−0.218
POST STOCK RETURN (+2,+20)	(−)	0.616	0.256	0.065
SHARE TURNOVER	(+)	3.423	< 0.001	0.526
MKTCAP	?	0.240	0.310	0.257
ANALYST	?	−0.038	0.422	−0.125
INSTITUTIONAL OWN	?	2.046	0.056	0.291
FINANCIAL	?	0.790	0.294	0.158
n		343[b]		
−2 log likelihood		338.79		
Model Chi-square		142.45		
p-value		< 0.001		
Pseudo R^2		34%		
Correctly Classified [baseline of no litigation is 81%]		91.0%		

---

[a] Two-tailed (one-tailed) tests are shown for variables without (with) a signed prediction.

[b] 23 firms were excluded from the regression due to missing market capitalization information and 15 observations were excluded due to missing *PRIOR STOCK RETURN*, *POST STOCK RETURN*, and *RESTATEMENT MAGNITUDE* data.

[c] To control for potential endogeneity between managers' disclosure choice (proxied for by *PROMINENCE INDEX*) and litigation likelihood, we run a first stage ordered probit model predicting *PROMINENCE INDEX* and generate three inverse Mills ratios. These ratios are included in the regression above. As their coefficients are not significant under any specification, and for the sake of brevity, we omit their coefficients from the table.

(*continued on next page*)

TABLE 7 (continued)

Variable Definitions:

LITIGATION = 1 if a class action lawsuit was filed against the firm in response to the restatement of their financial statements, and 0 otherwise;

CONCURRENT RETURN (−1, +1) = compounded raw return over the (−1, +1) interval with 0 being the date of the restatement announcement;

PRIOR STOCK RETURN (−252, −2) = compounded raw returns over one year, ending on the second day before the restatement announcement (−252, −2);

POST STOCK RETURN (+2, +20) = compounded raw return over the (+2, +20) interval with day 0 being the date of the restatement announcement; and

SHARE TURNOVER = probability that a share was traded within a given time period. It is calculated as: $[1 - \Pi(1 - \textit{volume traded}_t/\textit{total shares}_t)]$, accumulated over the one-year period ending on the second day prior to the restatement announcement date.

All other variables are defined in Table 4.

In Table 7, Panel B, we report results for the full model, after controlling for endogeneity using inverse Mills ratios.[15] Once again, the *PROMINENCE INDEX* coefficient is statistically significant. The only notable changes are that *RESTATEMENT MAGNITUDE*, *SEC*, and *FINANCIAL* are no longer significant.

Dismissed Lawsuits

During the discovery phase of litigation, extensive information about the nature of the accounting misstatement is revealed, resulting in some lawsuits being dismissed. To determine which lawsuits were dismissed, we used the Stanford Securities Class Action Clearinghouse website at the end of 2006 to review the status of each court case. In untabulated results, we find that 23 percent of the lawsuits are dismissed (17 of 73), and the rate of dismissal is virtually identical for firms providing high and medium prominence disclosure (24 percent and 23 percent, respectively). To provide direct evidence about the effect of dismissals on the likelihood of litigation, we estimate the logistic regression model with a reduced sample that omits dismissed lawsuits (untabulated). We find that the *PROMINENCE INDEX* coefficient remains significant. The other coefficients are substantively as reported, with the exception of the *SEC* and *INSTITUTIONAL OWN* coefficients, which are no longer statistically significant.

Based on this analysis, we reject null H2, which states that the extent of prominence given to an accounting restatement in the press release does not affect the likelihood of class action litigation. The overall evidence suggests that some of the companies providing medium prominence press release disclosure are rewarded with a lower likelihood of litigation. This result is robust to the treatment of dismissed lawsuits, so in contrast to the return differences, the reward appears to be permanent.

Robustness Tests

Because the average prominence score for revenue restatements is lower than the sample average (see Table 2, Panel C), we re-estimate the litigation model to determine if the litigation results are robust to the exclusion of those firms. We lose substantial power because the number of firms drops from 381 to 223 and the number of lawsuits drops from 73 to 37. We find that our results are weaker on this reduced sample, with the *PROMINENCE INDEX* coefficient no longer statistically significant (untabulated). We also estimate a model that includes *only* revenue restatement firms and find that *PROMINENCE INDEX* remains positive and significant (untabulated).

---

[15] In the first stage model that predicts *PROMINENCE INDEX* (untabulated), we find the number of management forecasts is significantly negatively associated with *PROMINENCE INDEX*. We also find significantly positive coefficients for *IRREGULARITY* and *FINANCIAL*, and significantly negative coefficients for *MAGNITUDE SOURCE*, *CONCURRENT RETURN*, *REVENUE*, *RULE CHANGE*, and *MKTCAP*. The coefficients for the three inverse Mills ratios in the second stage model are individually and jointly statistically insignificant, so we do not report them on Table 7.

Some prior studies of accounting restatements use information on the initiating entity (i.e., restating firm, auditor, or SEC) as a measure of severity (Palmrose et al. 2004; Efendi et al. 2007). To our knowledge, no prior study examines whether the identity of the restatement initiator is associated with the likelihood of litigation. We define an indicator variable that equals 0 for restatements initiated by the company, and 1 otherwise. We find that the coefficient on this variable is negative and significant. Thus, restatements initiated by an external party are less likely to result in litigation than those initiated by the firm. Coefficient signs and significance levels are unchanged for the other variables.

Finally, we remove firms with more than one restatement included in the GAO database and reach the same conclusions.

VII. CONCLUSION

We investigate whether the effect of prominence given to a restatement in the corporate press release affects market prices and litigation. We find that firms providing less prominent press release disclosure of a restatement are rewarded with a less negative return at the announcement, but much of this reward dissipates over the next 20 days. We also find that companies providing less prominent disclosure are less likely to be sued for securities fraud, and the decrease in litigation is a permanent reward.

Prior disclosure research focuses on managers' decision whether to issue a press release that warns of an earnings shortfall. We contribute to the disclosure literature by extending corporate disclosure into a new arena: press release prominence. Corporate practices in writing press releases are a promising avenue for future research because press releases are firms' primary timely disclosure medium, especially since Regulation FD now restricts selective communications. The findings in this study should be of interest to investor relations executives and other corporate managers with responsibility for press release disclosure. Regulators, especially the SEC, may want to issue guidelines to standardize press release disclosures so firms will not be rewarded for providing less prominent disclosure of an accounting restatement. Such guidelines should not simply require headline disclosure for all restatements, however, since this could increase the number of frivolous lawsuits.

APPENDIX
VARIABLE DEFINITIONS

Variable Name	**Variable Definition**
PROMINENCE INDEX	Disclosure emphasis is coded as 3 (high prominence) for any press release that mentions the earnings restatement in the headline. Disclosure quality is coded 2 (medium prominence) for press releases that do not mention the restatement in the headline, but that discuss the restatement in some detail within the body of the report. A code of 1 (low prominence) is assigned to press releases that only mention the restatement in the footnotes of the report.

(*continued on next page*)

Variable Name	**Variable Definition**
CAR	The cumulative abnormal return, calculated as the raw stock return minus the CRSP equally weighted market portfolio return measured over a short window (−1, +1) and a longer window (−1, +20), with day 0 indicating the restatement announcement.
LITIGATION	An indicator equal to 1 if a class action lawsuit was filed against the firm in response to the restatement of their financial statements, and 0 otherwise.
RESTATEMENT MAGNITUDE	The cumulative earnings effect of the restatement scaled by total assets measured as of the fiscal year-end prior to the restatement announcement, winsorized at the 5th and 95th percentiles. If the prior year's earnings were overstated (understated), then *MAGNITUDE* has a negative (positive) sign. When available, this data was collected from the initial press release announcing the restatement. When not available, it was collected from subsequent SEC filings.
MAGNITUDE SOURCE	An indicator variable equal to 1 if the restatement magnitude was not included in the press release initially announcing the restatement (i.e. the information was missing), and 0 otherwise. When missing from the press release, restatement magnitude information was pulled from SEC filings.
IRREGULARITY	An indicator equal to 1 if any of the following applies to the restatement: (1) variants of the words "fraud" or "irregularity" are used in the restatement announcement, (2) the firm is subject to an SEC Auditing and Accounting Enforcement Action, or (3) the firm initiates an independent investigation into the accounting issue and 0 otherwise. Independent investigations are those conducted by an (independent) audit committee, by a special committee of outside directors, by an outside forensic firm, or by an independent auditing firm who is not the usual auditor.
SHORT INTEREST	The percent of short interest [100 * (number of shares sold short)/(total shares outstanding)] as of the month prior to the restatement announcement.
SEC	An indicator equal to 1 if the press release indicates an SEC inquiry, and 0 otherwise.
REVENUE	An indicator equal to 1 if any part of the restatement is due to revenue recognition problems, and 0 otherwise.
RULE CHANGE	An indicator equal to 1 if a change in accounting rules was mentioned in the press release as the reason for the restatement, and 0 otherwise.
COUNT	The number of distinct accounting issues, as reported in the GAO database.
MKTCAP	The natural log of market capitalization (closing stock price * common shares outstanding) of the firm measured as of the end of the fiscal year prior to the restatement.
ANALYST	The number of analysts whose earnings forecasts are included in the consensus formed prior to the most recent quarterly or annual earnings announcement prior to the restatement announcement.
INSTITUTIONAL OWN	The level of institutional ownership at the end of the quarter prior to the restatement announcement, scaled by total shares outstanding on the same date.
FINANCIAL	An indicator equal to 1 if the firm operates in the financial services sector (SIC codes 6000–6999), and 0 otherwise.

(*continued on next page*)

Variable Name	**Variable Definition**
SURPRISE	The difference between actual quarterly earnings per share and the most recent consensus analyst earnings forecast issued prior to the earnings announcement, both reported by I/B/E/S, scaled by stock price measured just prior to the date of the forecast.
BEAT	An indicator equal to 1 if the firm beat analyst expectations (i.e. *SURPRISE* > 0), and 0 otherwise.
CONCURRENT RETURN	The compounded raw return over the (−1, +1) interval with 0 being the date of the restatement announcement.
PRIOR STOCK RETURN	The compounded raw returns over one year, ending on the second day before the restatement announcement (−252, −2).
POST STOCK RETURN	The compounded raw return over the (+2, +20) interval with day 0 being the date of the restatement announcement.
SHARE TURNOVER	A continuous variable measuring the probability that a share was traded within a given time period. It is calculated as: $[1 - \Pi(1 - \textit{volume traded}_t/\textit{total shares}_t)]$, accumulated over the one-year period ending on the second day prior to the restatement announcement date.
EXTERNAL INITIATOR	An indicator equal to 1 if an outside party initiated the restatement (e.g. auditor, SEC) as reported by the GAO database, and 0 otherwise.

REFERENCES

Allison, P. D. 2002. *Missing Data.* Series No. 07–136. Thousand Oaks, CA: Sage Publications.

Bowen, R. M., A. K. Davis, and D. A. Matsumoto. 2005. Emphasis on pro forma versus GAAP earnings in quarterly press releases: Determinants, SEC intervention, and market reactions. *The Accounting Review* 80 (4): 1011–1038.

Cox, J. D., and R. S. Thomas. 2006. Does the plaintiff matter? An empirical analysis of lead plaintiffs in securities class actions. *Columbia Law Review* 100: 101–155.

Efendi, J., M. R. Kinney, and E. P. Swanson. 2005. Can short sellers anticipate accounting restatements? Working paper, The University of Texas at Arlington and Texas A&M University.

———, A. Srivastava, and E. P. Swanson 2007. Why do corporate managers misstate financial statements? The role of in-the-money options and other incentives. *Journal of Financial Economics* 85 (3): 667–708.

Fama, E. F., and K. R. French. 1993. Common risk-factors in the returns on stocks and bonds. *Journal of Financial Economics* 33 (1): 3–56.

Feid, M. 2005. Back to the basics of press release writing. *Franchising World* 37 (9): 59–60.

Ferris, S. P., and A. C. Pritchard. 2001. Stock price reactions to securities fraud class actions under the private securities litigation reform act. Working paper, University of Michigan and University of Missouri–Columbia.

Field, L., M. Lowry, and S. Shu. 2005. Does disclosure deter or trigger litigation? *Journal of Accounting and Economics* 39 (3): 487–507.

Francis, J., D. Philbrick, and K. Schipper. 1994. Shareholder litigation and corporate disclosures. *Journal of Accounting Research* 32 (2): 137–164.

Gande, A., and C. M. Lewis. 2009. Shareholder initiated class action lawsuits: Shareholder wealth effects and industry spillovers. *Journal of Financial and Quantitative Analysis* (44) 4, doi: 10.1017/S0022109009990202.

General Accounting Office (GAO). 2002. *Financial Statement Restatements: Trends, Market Impacts, Regulatory Responses, and Remaining Challenges Report 03–138.* Washington D.C.: Government Printing Office.

———. 2003. *Financial Statement Restatements Database.* Report 03–395R. Washington D.C.: Government Printing Office.

Glass Lewis & Co LLC. 2006. *Getting It Wrong the First Time.* A look at 2005's record-breaking year for corporate restatements shows why investors can't afford a return to pre-Enron securities regulation. San Francisco, CA: Glass Lewis.

Gordon, E. A., E. Henry, M. Peytcheva, and L. Sun. 2008. Disclosure credibility and market reactions to restatements. Working paper, Temple University, University of Miami, Lehigh University, and Rutgers University.

Graham, J. R., C. R. Harvey, and S. Rajgopal. 2005. The economic implications of corporate financial reporting. *Journal of Accounting and Economics* 40: 3–73.

Grundfest, J. A., and M. A. Perino. 1997. *Ten Things We Know and Ten Things We Don't Know About the Private Securities Litigation Reform Act of 1995.* Joint written testimony before the Subcommittee on Securities of the Committee on Banking, Housing, and Urban Affairs, United States Senate, on July 24, 1997. Washington, D.C.: Government Printing Office.

Heckman, J. 1979. Sample selection bias as a specification error. *Econometrica* 47: 153–161.

Hennes, K., A. Leone, and B. Miller. 2008. The importance of distinguishing errors from irregularities in restatement research: The case of restatements and CEO/CFO turnover. *The Accounting Review* 83 (6): 1487–1519.

Hirshleifer, D., and S. H. Teoh. 2003. Limited attention, information disclosure, and financial reporting. *Journal of Accounting and Economics* 36: 337–386.

———, S. S. Lim, and S. H. Teoh. 2004. Disclosure to an audience with limited attention. Working paper, University of California, Irvine and DePaul University.

Jones, C. L., and S. E. Weingram. 1996. The determinants of 10b-5 litigation risk. Working paper, Stanford Law School and The George Washington University.

Kasznik, R., and B. Lev. 1995. To warn or not to warn: Management disclosures in the face of an earnings surprise. *The Accounting Review* 70 (1): 113–134.

Kothari, S. P., S. Shu, and P. Wysocki. 2008. Do managers withhold bad news? *Journal of Accounting Research* 47 (1): 241–276.

Lev, B., S. Ryan, and M. Wu. 2008. Rewriting earnings history. *Review of Accounting Studies* 13 (4): 419–451.

Mahoney, W. F. 1991. *Investor Relations: The Professional's Guide to Financial Marketing and Communications.* New York, NY: New York Institute of Finance.

Myers, L., S. Scholz, and N. Sharp. 2008. Restating under the radar: Determinants of restatement disclosure choices and the related market reactions. Working paper, University of Arkansas, University of Kansas, and Texas A&M University.

Palmrose, Z-V., V. J. Richardson, and S. Scholz. 2004. Determinants of market reactions to restatement announcements. *Journal of Accounting and Economics* 37: 59–89.

———, and S. Scholz. 2004. The circumstances and legal consequences of non-GAAP reporting: Evidence from restatements. *Contemporary Accounting Research* 21 (1): 139–180.

PricewaterhouseCoopers LLP 2003. Securities litigation study 6. Available at: *www.10b5.com/2003_study.pdf.*

Reilly, D. 2006a. Sarbanes-Oxley changes take root. *The Wall Street Journal* (March 3).

———. 2006b. No more "stealth restating." *The Wall Street Journal* (September 21): C1, C3.

Rogers, J. L., and A. Van Buskirk. 2008. Shareholder litigation and changes in disclosure behavior. *Journal of Accounting and Economics* 47 (1–2): 136–156.

Skinner, D. 1994. Why firms voluntarily disclose bad news. *Journal of Accounting Research* 32 (1): 38–60.

———. 1997. Earnings disclosures and stockholder lawsuits. *Journal of Accounting and Economics* 23: 249–282.

Tucker, J. W. 2007. Is openness penalized? Stock returns around earnings warnings. *The Accounting Review* 82 (4): 1055–1087.

Wu, M. 2003. Earnings restatements: A capital market perspective. Working paper, Hong Kong University of Science and Technology.

THE ACCOUNTING REVIEW
Vol. 84, No. 5
2009
pp. 1521–1552

American Accounting Association
DOI: 10.2308/accr.2009.84.5.1521

Big 4 Office Size and Audit Quality

Jere R. Francis
University of Missouri–Columbia

Michael D. Yu
Washington State University

ABSTRACT: Larger offices of Big 4 auditors are predicted to have higher quality audits for SEC registrants due to greater in-house experience in administering such audits. We test this prediction by examining a sample of 6,568 U.S. firm-year observations for the period 2003–2005 and audited by 285 unique Big 4 offices. Results are consistent with larger offices providing higher quality audits. Specifically, larger offices are more likely to issue going-concern audit reports, and clients in larger offices evidence less aggressive earnings management behavior. These findings are robust to extensive controls for client risk factors and to controls for other auditor characteristics. While the evidence suggests audit quality is higher on average in larger Big 4 offices, we make no claims that audit quality is unacceptably low in smaller offices.

Keywords: *audit quality; Big 4 accounting firms; earnings quality; accruals; earnings benchmarks; going-concern audit reports.*

Data Availability: *Data used in this study are available from public sources identified in the paper.*

I. INTRODUCTION

This study extends recent research analyzing the effects of client influence and auditor industry expertise in individual practice offices of Big 4 accounting firms (Reynolds and Francis 2000; Craswell et al. 2002; Ferguson et al. 2003), and investigates a fundamental question that has not been addressed in prior studies: Is Big 4 audit quality uniform across small and large practice offices? Our prediction is that audits are of higher quality in larger Big 4 offices because auditors in these offices have more collective experience in administering the audits of public companies (SEC registrants). Thus, a large

---

We thank the editors, Steve Kachelmeier and Dan Dhaliwal, and the two anonymous referees for their many constructive suggestions. We also appreciate feedback on earlier versions of the study presented at the 2007 American Accounting Association Annual Meeting, the 2007 European Audit Research Network Symposium, and workshops at University of Auckland, Bond University, University of Colorado, Indiana University, University of Missouri–Columbia, University of Melbourne, Tilburg University, Washington State University, and Yale University, and especially the comments of Paul Brockman, Inder Khurana, Elaine Mauldin, Raynolde Pereira, Kenny Reynolds, Phil Shane, Stephen Taylor, and Marlene Willekens. This study is supported by a grant from the *PwC INQuires* research program of PricewaterhouseCoopers.

Editor's note: Accepted by Steven Kachelmeier, with thanks to Dan Dhaliwal for serving as editor on a previous version.

Submitted: August 2007
Accepted: November 2008
Published Online: September 2009

office will have greater in-house expertise in detecting material problems in the financial statements of SEC clients. By implication, auditors in smaller Big 4 offices have less experience and therefore develop less skill in detecting such problems.

To test the relation between Big 4 office size and audit quality, we examine the association of office size with going-concern audit reports and client earnings properties (abnormal accruals and earnings benchmark tests). Big 4 office size is measured by the fees received from SEC registrants, and the results are robust to alternative measures using total audit fees (audit plus nonaudit), and ranks of fees. Importantly, the models include extensive controls to assure that office size is not capturing the effects of omitted client risk factors or auditor characteristics such as tenure and industry expertise, although we cannot entirely rule this out. The models are also estimated as fixed- or random-effect models as an additional control for omitted variables.

We find that larger offices are more likely to issue going-concern reports, and that their going-concern reports are more accurate in terms of predicting next-period client bankruptcy. Clients audited by larger offices are also less likely to have aggressively managed earnings as evidenced by smaller abnormal accruals and a lower likelihood of meeting benchmark earnings targets (small profits and small earnings increases). Overall, these results reinforce the importance of the local office unit of analysis in audit research and show there is significant variation in audit outcomes across Big 4 offices, with the evidence consistent with the premise that larger offices provide higher quality audits. As reported in Section V, however, the results are less robust when office size is based on the number of SEC clients rather than total office fees.[1]

The study is subject to the following caveats. First, our evidence does not indicate small offices fail to meet minimum standards of audit quality; however, the findings do point to systematically higher quality by larger offices relative to smaller offices of Big 4 accounting firms. Second, the analysis is based on public company (SEC) clienteles, and hence the knowledge and expertise analyzed in the study is an office's expertise in dealing with SEC registrants. The analysis of private company clienteles is beyond the scope of this study, and cannot be undertaken due to the lack of publicly available data. A third caveat is that audits are wholly attributed to the engagement office of record based on the audit report filed with the SEC. We recognize that multiple offices of a Big 4 firm may contribute to an audit engagement, although this is not determinable with publicly available data. However, the engagement office that contracts with the client has primary responsibility for the audit, including overseeing work performed by other offices. Thus, the engagement office's audit team makes critical judgments on audits, and of course the engagement partner issues the final audit report on engagement office letterhead. Therefore, it is reasonable to attribute the audit entirely to the engagement office for the purpose of our study, even though other offices may participate in the audit (albeit with oversight by the engagement office). In a practical sense, the extent to which small offices participate on audit engagements of large offices (and *vice versa*) would only neutralize office size differences and, therefore, should work against the predicted office size/audit quality relation.

The next section develops the study's hypothesis and explains why an office-level analysis is important. Section III presents the research design, sample selection, and descriptive statistics. Section IV discusses the primary empirical results, and Section V reports sensitivity tests and robustness checks. Section VI concludes the study.

---

[1] A concurrent study by Choi et al. (2007) uses a different design and sample, but also reports a negative association between office size and absolute abnormal accruals.

II. BACKGROUND AND HYPOTHESIS DEVELOPMENT

Wallman (1996) and Francis, Stokes and Anderson (1999) argue that local practice offices are the primary decision-making unit within Big 4 auditing firms and, therefore, an important unit of analysis in audit research. Big 4 firms have decentralized organizations and operate through a network of semi-autonomous practice offices. Local offices contract with clients, administer audit engagements, and issue audit reports signed on the local office letterhead. Accounting professionals are typically based in specific practice offices and audit clients in the same geographic locale; hence, their expertise and knowledge is both office- and client-specific (Francis, Stokes, and Andersen 1999; Ferguson et al. 2003). This decentralized office structure reduces information asymmetry and enables Big 4 auditors to develop better knowledge of existing and potential clients in a particular location. Clients, in turn, have greater knowledge of and confidence in the expertise of locally based personnel who actually perform audits (Carcello et al. 1992). The above argument assumes that Big 4 firms are unable to fully achieve uniform audit quality across offices, and that a certain amount of overall audit expertise is office-specific (Francis et al. 2005; Vera-Muñoz et al. 2006).

Given the above discussion, our argument is that a large office has more "in-house" experience in dealing with public companies (SEC registrants) and, hence, more collective human capital in the office. Experience is an important dimension of human capital (Becker 1993), and a larger office with more engagement hours therefore provides its auditors with greater opportunities to acquire expertise in detecting material problems in the financial statements of SEC registrants. As a consequence, auditors in larger offices are more likely to detect and report material problems in the financial statements, or require clients to correct the statements before issuance.[2]

Auditors working in a large office will have more peers with whom to consult and, hence, have a better local support network. Danos et al. (1989) report that auditors are most likely to consult their peers within the same office when problems arise rather than broader consultation with colleagues in other offices or the national office. It follows that larger offices also have the potential to produce higher quality audits because of their greater in-house networking/consultation opportunities. We acknowledge that in the post-SOX era there may be more firm-wide consultations to facilitate better quality audits and, to the extent this is the case, it would work against the predicted office-size effect.

Based on the above discussion, we believe audit quality is not uniform across Big 4 offices, and the study's hypothesis in alternative form is:

> Larger offices of Big 4 accounting firms provide higher quality audits, where higher quality audits are inferred by the auditor's likelihood of issuing a going-concern audit report (and accuracy of the report in predicting client bankruptcy), and the degree to which clients evidence earnings management behavior.

---

[2] A secondary argument is that larger offices also have deeper reserves of personnel (slack) to mitigate the effects of high employee turnover in the public accounting industry. Satava (2003) reports that the large national accounting firms have a turnover rate of around 25 percent or the loss of one in four employees annually. Auditor turnover results in the loss of auditor expertise and knowledge, and especially the specific knowledge between an auditor and a client. However, because a large office has a bigger pool of employees, it is better able to replace audit team members with experienced auditors. The same logic applies to the mandatory rotation of engagement partners and concurring review partners. A larger office has a deeper reserve of partner expertise to draw on when mandatory rotation occurs and new partners must be assigned to clients. Hence, in a large office, there is more likely to be continuity in the office's expertise in administering SEC audit engagements from one period to another and from one audit team to another.

The null hypothesis is that audit quality is uniform across office size. While we do not expect this to be the case (and the evidence does not support that it is), it cannot be ruled out *a priori*. The Big 4 firms are organized as national partnerships with national administrative offices that set firm-wide policies and provide technical support for their city-based practice offices. Under this alternative view of the audit firm, audit expertise and knowledge can be captured by the firm as a whole and distributed *uniformly* across offices. This view is supported by the fact that the Big 4 firms have national training programs, standardized audit programs, and firm-wide knowledge-sharing practices supported by information technology. Auditors travel, to some extent, between offices, and may also be reassigned to other offices, both of which can spread expertise across offices. However, Vera-Muñoz et al. (2006) point out that firm-wide knowledge sharing has practical limitations, and for this reason it is an open empirical question as to what extent these firm-wide mechanisms can effectively mitigate the hypothesized office-size effect on Big 4 audit quality. Recent changes implemented by Sarbanes-Oxley, such as the annual inspections undertaken by the PCAOB, have created additional incentives for accounting firms to strengthen their internal procedures to ensure uniform audit quality across practice offices. To the extent that accounting firms have restructured their operations to improve and standardize firm-wide audit quality, this would work against finding the hypothesized office-size effect.

III. RESEARCH DESIGN

Audit quality is inferred by examining client earnings properties and implied earnings management behavior with respect to abnormal accruals and earnings benchmark targets (e.g., Becker et al. 1998; Frankel et al. 2002). Earnings management per se does not violate generally accepted accounting principles. However, firms that manage earnings are viewed as having lower quality earnings (e.g., Frankel et al. 2002), and Levitt (1998) suggests that aggressive earnings management can result in materially misleading financial reports.

We test if client earnings metrics differ across office sizes of Big 4 firms. Specifically, we expect clients in larger offices will evidence less earnings management behavior (smaller abnormal accruals and less likelihood of meeting benchmark earnings targets) after controlling for client risk factors. The reason is that auditors in larger offices are expected to have more expertise in detecting and deterring aggressive earnings management behavior (Francis, Maydew and Sparks 1999). In addition, we test if an auditor's propensity to issue a going-concern report (and the accuracy of going-concern reports in predicting client bankruptcy) is increasing in office size. Again, the conjecture is that auditors in larger offices have more expertise in identifying the circumstances that warrant a going-concern report. The auditor's likelihood of issuing going-concern audit reports has been used in prior research to test for differential audit quality (Reynolds and Francis 2000; Craswell et al. 2002; DeFond et al. 2002).

The specific office administering an audit engagement is identified from the letterhead of the audit report filed with the SEC, as reported in Audit Analytics. We use an office's aggregate audit fees each year to measure office size using all observations in the Audit Analytics database with fee data. Audit fees are directly related to engagement hours, and offices with higher fees will therefore have more hours of experience in the audits of SEC registrants. The log of office fees (denoted *lnOFFICE*) is the functional form used in the multivariate analyses due to skewness in the distribution of office-level audit fees (see Table 1). Results are robust to using total fees (audit and nonaudit) to measure office size as in Craswell et al. (2002), and to using log of ranks where the 805 offices in the sample are rank-ordered from 1 to 805 based on their audit fees.

TABLE 1
Big 4 Accounting Firm Office Size Based on Pooled 2003–2005 Data in Audit Analytics[a]

Panel A: Auditor Office Size for 805 Office Years in the Study Based on Audit Fee Revenues (in $ millions)[b]

Auditor	# of Office Years	Mean	Median	Std. Dev.	Min.	Q1	Q3	Max.	Aggregate Fees
Deloitte & Touche LLP	186	27.69	12.325	49.85	0.17	2.89	35.14	439.93	5,149
Ernst & Young LLP	219	25.76	13.45	36.88	0.15	5.76	29.41	270.28	5,641
KPMG LLP	220	20.97	7.335	36.76	0.2	3.55	21.65	293.24	4,614
PricewaterhouseCoopers LLP	180	42.19	14.945	79.96	0.06	5.63	45.185	623.53	7,594

Panel B: Auditor Office Size for 805 Office Years in the Study Based on Number of Clients[c]

Auditor	# of Office Years	Mean	Median	Std. Dev.	Min.	Q1	Q3	Max.	Aggregate Clients
Deloitte & Touche LLP	186	36.79	15	82.76	1	5	42	764	6,843
Ernst & Young LLP	219	33.6	15	58.48	1	8	30	352	7,359
KPMG LLP	220	24	11	38.93	1	5	27	321	5,280
PricewaterhouseCoopers LLP	180	42.47	16	78.22	1	7	31.5	503	7,645

---

[a] This table provides office size descriptive statistics for the 285 unique U.S. Big 4 offices in the study. The 285 offices are distributed as follows: Deloitte (65), Ernst & Young (78), KPMG (77), and PricewaterhouseCoopers (65). Each office can appear up to three times over the three-year sample period (2003–2005), and there are 805 "office years" in total. The data in this table are based on 27,127 firm-year (client) observations of Big 4 auditors and represent *all* Audit Analytics observations located in the U.S. with audit fee data for fiscal years 2003 through 2005, using the Compustat year convention.

[b] Panel A reports summary statistics based on office-level audit fee revenues (in $ millions), per office-year.

[c] Panel B reports summary statistics based on the number of clients audited by each office, per office-year.

Up to this point, the auditor's incentives have not been explicitly considered in the analysis. Prior research argues that auditors may acquiesce and report favorably in order to retain influential clients, particularly if a client is large relative to the size of the engagement office (Reynolds and Francis 2000). The proposition that auditors report favorably to retain important clients is known as economic bonding (DeAngelo 1981). Following Craswell et al. (2002), we measure the auditor's incentives with respect to a client's influence (*INFLUENCE*) on the local office as the ratio of the client's fees for all services, to the sum of fees for all clients of the engagement office for a given year. While economic bonding implies the impairment of auditor independence, Reynolds and Francis (2000) report evidence of the opposite, namely, that Big 4 auditors report more conservatively for larger influential clients in engagement offices. The explanation in Reynolds and Francis (2000) is that the auditor's incentive to avoid costly litigation from misreporting by important clients is stronger than the incentive to acquiesce and report favorably. We make no directional prediction in this study, but include *INFLUENCE* to control for the auditor's office-level incentives with respect to influential clients. It is also important to note that *INFLUENCE* measures a client's size ***relative*** to an office, and that it is distinctly different from both ***absolute*** client size and ***absolute*** office size. A client of a given absolute size (e.g., fees of $1 million) could be a relatively large or small client depending on the absolute size of the engagement office.

Two other important auditor characteristics are controlled for in all models to assure that the results for office size are not the consequence of correlated omitted auditor variables. First, we control for auditor tenure because Johnson et al. (2002) find that short auditor tenure is associated with lower client earnings quality. Following Johnson et al. (2002), we include the variable *TENURE*, which is coded 1 if tenure is three years or less, and 0 otherwise, to assure that office size is not in some way confounded by systematic differences in auditor tenure across practice offices. Second, we control for the auditor's industry expertise to assure that office size is not capturing an omitted variable with respect to the auditor's industry expertise. Prior studies argue that industry expertise increases audit quality (Balsam et al. 2003), and two-digit SIC codes are used to calculate industry-expertise measures at the national level (based on all clients of auditors) and office level (based on city-specific clienteles of auditors). As in Francis et al. (2005), national industry expertise is an indicator variable that is coded 1 if the auditor is the national audit fee leader (*NATIONAL-LEADER*), and an office is classified as an industry expert if it is the city-specific industry fee leader (*CITY-LEADER*). The results are unchanged if an auditor's actual national- and city-level market shares of fees are used in lieu of indicator variables for industry leadership.

The final two control variables used in all models are the number of client operating segments (*OPSEG*) and number of geographic segments (*GEOSEG*) as reported in Compustat. If no segment data is reported in Compustat for a given observation, then we assign a value of 1. The intuition for these two variables is that clients with multiple operating divisions or geographical segments are more likely to require the use of additional offices to assist the lead engagement office in completing the audit. Thus, the purpose of the segment variables is to give confidence that test results for the size of the primary engagement office are robust to control for the potential confounding effects of other offices that may participate in an audit engagement. We make no prediction on the sign, although it is possible that the participation of multiple offices increases audit quality since there would be more offices involved in the engagement.[3]

Accruals Quality

The first dependent variable is abnormal accruals (Jones 1991). A large abnormal or discretionary component of accruals is indirect evidence of earnings management behavior and lower earnings quality. Kothari et al. (2005) argue that the discretionary accruals model might be misspecified when applied to samples of firms with extreme performance, and suggest that controlling for current firm performance will increase the power of the Jones model. We use ordinary least-squares (OLS) to estimate the following performance-adjusted Jones model for the full Compustat sample by fiscal year and two-digit industry SIC code (with a minimum of ten observations required for an industry to be included in a year), and controlling for concurrent firm performance with *NI*:

$$TA = \alpha + \beta_1 \Delta REV + \beta_2 PPE + \beta_3 NI + \varepsilon \tag{1}$$

where *TA* is total accruals; ΔREV is revenues in year t less revenues in year $t-1$; *PPE* is gross property, plant, and equipment; and *NI* is operating income after depreciation.[4] All variables are deflated by lagged total assets. The absolute value of residuals from Equation (1) is used to measure discretionary accruals since individual firms may have incentives to manage earnings either up or down depending on particular circumstances (Warfield et al. 1995). However, it has been argued that auditors are more concerned with constraining income-increasing accruals (Becker et al. 1998). Therefore, as an additional analysis, "signed" accruals are also examined by partitioning observations into those with income-increasing and income-decreasing abnormal accruals.

We use the following model adapted from Reynolds and Francis (2000) to test the relation between accruals and office size:

$$ACCRUALS = \lambda_0 + \lambda_1 lnOFFICE + \mathbf{X}'\boldsymbol{\lambda} + \ddot{\upsilon}. \tag{2}$$

OLS is used to estimate Equation (2), and we follow Newey and West (1987) to correct for heteroscedasticity and first-order autocorrelation (serial dependence). Results are robust to alternative estimations using firm fixed-effect models (to control for omitted variables) and linear mixed models with multilevel random effects.

---

[3] *OPSEG* and *GEOSEG* are insignificant in most of the tests. While these variables control for the effect of multiple operating and geographic segments on the study's dependent variables, they do not directly test if the office-size effects are systematically different for firms with single segments versus multiple segments. To directly analyze this question we re-code *OPSEG* and *GEOSEG* as equal to 1 if a firm has a single operating or geographic segment, respectively; otherwise, the segments variables are re-coded to 0. We then use the re-coded segment variables in the models, along with the interaction of each segment variable with the test variable *lnOFFICE*. Results of this new model specification are as follows. In the accruals tests and the two benchmark earnings tests, *lnOFFICE* is significant at the 0.05 level or less in all tests, and the interaction terms are not significant at the 0.10 level, indicating that the results for office size are consistent for firms with single operating/ geographic segments and firms with multiple segments. For the going-concern tests, the results indicate that auditors in larger offices are more likely to issue going-concern reports, and this result is even stronger for clients with a single operating or geographic segment.

[4] We use operating income after depreciation (Compustat data item 178) as a performance control because it excludes nonoperating income, special items, and other items that are of a more discretionary nature. Kothari et al. (2005) use income before extraordinary items (Compustat data item 18), and our results are robust to this alternative definition of income. As a further sensitivity we also test net income (Compustat data item 172), although this measure might be noisier since it includes both extraordinary and nonoperating items. When using net income the results are comparable for absolute and negative abnormal accruals, but positive abnormal accruals are not significant at the 0.10 level.

The dependent variable in Equation (2) is abnormal accruals (*ACCRUALS*) and is the residual of Equation (1) above. The test variable is office size (*lnOFFICE*) and is defined as the log of total office-specific audit fees of all clients per fiscal year. Since larger values of abnormal accruals imply more client discretion and lower earnings quality, we expect the coefficient on office size will be negative if auditors in larger offices allow their clients less discretion over the use accruals to manage earnings.

X is a vector of control variables that includes *INFLUENCE*, *TENURE*, *NATIONAL-LEADER*, *CITY-LEADER*, *OPSEG*, and *GEOSEG*, which were discussed in the previous section. The remaining control variables represent an extensive set of client variables used in prior research, plus other variables to assure the effects of office size are not the consequence of omitted client risk factors. Becker et al. (1998) find that larger clients are more likely to have higher earnings quality, so we expect that absolute client size (*SIZE*), measured as log of total assets ($ millions), will be negatively correlated with accruals. Menon and Williams (2004) find that sales growth (*SALESGROWTH*) is positively associated with abnormal accruals and we include the one-year growth rate in sales as a control. Based on the analysis in Hribar and Nichols (2007), we also control for the volatility of sales growth (*SALESVOLATILITY*) measured as the standard deviation of sales for the most recent three fiscal years. Dechow et al. (1995) show that operating cash flows (*CFO*) influence the magnitude of discretionary accruals, and we expect that higher operating cash flows are associated with lower discretionary accruals. In addition, Doyle et al. (2007) and Hribar and Nichols (2007) report a positive association between cash flow volatility and accruals, so we include volatility (*CFOVOLATILTY*) measured as the standard deviation of cash flows for the most recent three fiscal years. Doyle et al. (2007) find an association between internal control deficiencies (reported under Sarbanes-Oxley) and the contemporaneous quality of a firm's earnings. To control for this we use the variable *WEAKNESS*, which is the number of material internal control weaknesses in a fiscal year as reported in the Audit Analytics database. The variable is coded 0 if an observation has no deficiencies reported in Audit Analytics.

Three variables are included in the model to control for the affects of debt and financial distress: *DEBT*, *LOSS*, and *BANKRUPTCY*. DeFond and Jiambalvo (1994) argue that companies with more debt (*DEBT*) have greater incentives to use accruals to increase earnings due to debt covenant constraints, and predict that debt level should be positively correlated with discretionary accruals. Firms with negative earnings (*LOSS*) are also expected to have a negative association with accruals quality. The intuition is that firms that report losses have lower incentives to manage discretionary accruals than do firms that report positive earnings. As in Reynolds and Francis (2000), a summary measure of financial distress is also used based on the Altman bankruptcy model (*BANKRUPTCY*). Lower values indicate more financial distress so that a negative association is expected with accruals.[5] Following Matsumoto (2002) and Hribar and Nichols (2007), we include two market-based variables to control for market incentives: stock return volatility (*VOLATILITY*) and market-to-book ratio (*MB*), which is a proxy for risk and growth. Inclusion of these market-based variables is motivated by the fact that capital market pressure can influence earnings management behavior. Riskier firms and growth firms may have greater incentives to manage earnings in order to meet market expectations, and we expect both variables to be positively correlated with accruals.

---

[5] The following equation from Altman (1983) is used to calculate this measure: 0.717 * working capital/total assets + 0.847 * retained earnings/total assets + 3.107 * earnings before interest and taxes/total assets + 0.42 * book value of equity/total liabilities + 0.998 * sales/total assets.

Benchmark Earnings Targets

Earnings distributions have been used to test earnings quality and earnings management behavior. Prior studies conclude that firms are systematically managing earnings to meet benchmark targets because there are an abnormally high proportion of firms that just "meet or beat" benchmarks and an abnormally low proportion of firms just below benchmark targets (Burgstahler and Dichev 1997; Degeorge et al. 1999). We use a probit model to test two common benchmarks: reporting small positive profits (avoiding losses), and reporting small positive earnings increases (avoiding earnings declines). Earnings are assumed to be of higher quality (less subject to earnings management) if a firm does not systematically meet benchmark earnings targets. The prediction is auditors in larger offices are more likely to detect and constrain aggressive earnings management and that clients in larger offices are therefore less likely to meet benchmark targets.

A probit model is estimated for the pooled sample with clustered robust standard errors to correct for heteroscedasticity and serial dependence (Rogers 1993):

$$PROBIT[BENCHMARK = 1] = f(\beta_0 + \beta_1 lnOFFICE + \mathbf{X}'\boldsymbol{\beta} + \varepsilon) \quad (3)$$

where *BENCHMARK* is coded as 1 if a firm reports small positive earnings (or small earnings increase), and 0 otherwise. As a sensitivity analysis we also estimate a random effect probit model and the results are consistent with the model in Equation (3).[6] **X** is a vector of control variables that is the same as those in Equation (2) for abnormal accruals.

To test the reporting of small profits, we classify a client as reporting small positive earnings if its net income deflated by lagged total assets is between 0 and 5 percent. Frankel et al. (2002) and Carey and Simnett (2006) use a cutoff value of 2 percent, and our results are robust to this smaller cutoff level, as well as intermediate cutoffs of 3 and 4 percent. To test small earnings increases, we classify a client as reporting a small earnings increase if the change in its net income deflated by lagged total assets is between 0 and 1.3 percent. Frankel et al. (2002), Ashbaugh et al. (2003), and Carey and Simnett (2006) use a slightly larger cutoff value of 2.0 percent, and our results are robust to this larger cutoff level, as well as cutoffs 1 and 1.5 percent.

Going-Concern Audit Reports

A probit model adapted from prior studies tests if the propensity to issue going-concern audit reports differs across office size (e.g., Reynolds and Francis 2000; Craswell et al. 2002; DeFond et al. 2002). If larger offices have more expertise, then they should be better able to identify going-concern problems and issue more timely going-concern reports. Hence, we predict that office size is positively associated with the probability of issuing going-concern reports. The following probit model is estimated for the pooled sample with clustered robust standard errors to correct for heteroscedasticity and serial dependence (Rogers 1993):

$$PROBIT[GCREPORT = 1] = f(\beta_0 + \beta_1 lnOFFICE + \mathbf{X}'\boldsymbol{\beta} + \varepsilon) \quad (4)$$

---

[6] The standard procedure for cross-sectional panel data is to estimate a random-effect probit model that corrects for serial correlation as well as controls for omitted firm-level variables (Wooldridge 2002).

where *GCREPORT* is a dichotomous variable that takes the value of 1 if a client receives a going-concern audit report, and 0 otherwise. The test variable is *lnOFFICE*, and **X** is a vector of control variables that includes *INFLUENCE*, *TENURE*, *NATIONAL-LEADER*, *CITY-LEADER*, *OPSEG*, and *GEOSEG* as in the accruals and earnings benchmark tests. Predicted signs on these control variables are opposite that in Equations (2) and (3) because a larger value of the dependent variable denotes higher quality audits.

We also control for client risk factors that have specifically been shown in prior research to explain going-concern opinion reporting (Reynolds and Francis 2000; DeFond et al. 2002). The additional variables with the expected signs in parenthesis are *SALESVOLATILITY* (+), *SIZE* (−), *CASH* (−), *PRIORGC* (+), *REPORTLAG* (+), *DEBT* (+), *LOSS* (+), *LAG LOSS* (+), *BANKRUPTCY* (−), *LAG RETURN* (−), *VOLATILITY* (+), and *MB* (+). *SALESVOLATILITY* is the standard deviation of the last three years' sales and is expected to have a positive association with going-concern reports due to higher operating risk. *SIZE* is log of total assets of the client, and is expected to be negatively correlated with the dependent variable because larger clients have more resources to stave off bankruptcy and therefore are less likely to fail. *CASH* is a liquidity measure that is the sum of the firm's cash and investment securities, scaled by total assets. A firm with more liquid assets has the resources to deal with financial difficulties; this variable is expected to be negatively associated with the probability of a going-concern opinion. We include a dummy variable *PRIORGC*, which takes the value of 1 if a company received a going-concern opinion in the previous period as companies are more likely to receive a going-concern report if they received a prior-year going-concern qualification (Reynolds and Francis 2000).[7] *REPORTLAG* is a timeliness variable measuring the number of days between the fiscal year-end and the earnings announcement date. Prior research finds that going-concern opinions are associated with longer reporting delays (Raghunandan and Rama 1995; Carcello et al. 1995; DeFond et al. 2002). *DEBT* is total liabilities deflated by total assets, and *LOSS* is a dummy variable that takes the value of 1 if the company has an operating loss in the current year. High-debt firms and firms reporting losses are more likely to fail and therefore more likely to receive going-concern reports. *BANKRUPTCY* is the Altman Z-score (Altman 1983), which measures the probability of bankruptcy. The market measures *VOLATILITY* and *MB* are positively associated with going-concern reports because riskier growth firms are more likely to fail, while firms with higher returns in the prior year (*LAG RETURN*) are more likely to be performing well and are less likely to fail.

Sample Selection

The sample covers the three-year period 2003 through 2005 based on Compustat year definitions. The Big 4 auditors are Deloitte, Ernst & Young, KPMG, and PricewaterhouseCoopers, and office size each year is based on aggregate yearly audit fees for each office in Audit Analytics. The engagement office is determined from the audit report letterhead in SEC filings as reported in Audit Analytics, and the full population of observations with audit fee data is used for the calculation of the fee-based measure of office size (before merging with Compustat). Auditor industry leadership is also based on the full Audit Analytics population with fee data. An audit firm (office) is denoted the national (city-specific) industry leader if it has the largest client audit fees for a specific industry (city-specific industry) in a fiscal year. Industry leadership is based on two-digit SIC classification, which is also reported in Audit Analytics.

---

[7] An alternative design is to examine first-time going-concern reports. There are 173 going-concern reports in our sample including 78 first-time reports. If we restrict the analysis to first-time reports, the power of the test is reduced due to small sample size and office size is significant at the 0.11 level (one-tailed).

After merging the Audit Analytics sample with Compustat, we exclude non-Big 4 auditors and observations in Compustat with missing financial data. In addition, the financial sector (SIC codes 60–69) and regulated industries (SIC codes 44–49) are excluded because their accruals structure is qualitatively different. The final sample consists of 6,568 firm-year observations and 2,572 unique firms for fiscal years 2003 through 2005 with the required auditor office-level data from Audit Analytics and Compustat financial data.

The 6,568 firm-years in the sample are audited by 285 unique Big 4 offices, which are distributed as follows: PricewaterhouseCoopers (65), Deloitte (65), Ernst & Young (78), and KPMG (77). Each office can appear up to three times (2003, 2004, and 2005) and there are a total of 805 "office-year" observations in the pooled sample. Note that a few offices do not appear in all three years, and the results are robust to excluding these offices from the study. Table 1 reports office size descriptive statistics for the 805 office-year observations using the full Audit Analytics population with audit fee data (before merging with Compustat). The smallest office of each Big 4 firm has a single SEC registrant and less than $1 million in audit fees. In contrast, the largest office of each Big 4 firm has over 300 clients and audit fees in excess of $250 million. The median number of clients per office ranges from 11 for KPMG to 16 for PricewaterhouseCoopers, and median fees per office range from $7 million for KPMG to $15 million for PricewaterhouseCoopers.

Descriptive Statistics

Table 2 reports summary statistics of variables used in the study. The variable for office size (*OFFICE*) has mean (median) audit fees of $65.4 ($40.8) million.[8] The log of *OFFICE* (denoted *lnOFFICE*) is the functional form used for the multivariate tests due to skewness and some extreme large values in the distribution of office-level fees.

Signed abnormal accruals (*ACCRUALS*) are winsorized at −0.999 and +0.999, and both the mean and median are 0.009. Distribution of the absolute value of abnormal accruals (*ABS_ACCRUALS*) has a mean of 0.258 and median of 0.124, indicating some larger right-tailed values. For the benchmark earnings tests, a client reports small positive earnings (*SMALL_PROFIT*) if its net income deflated by lagged total assets is between 0 and 5 percent. To test small earnings increases, we classify a client as reporting a small earnings increase (*SMALL_INCREASE*) if its change in net income deflated by lagged total assets is between 0 and 1.3 percent. A total of 23.3 percent of observations report small profits over the sample period, and 12.9 percent of observations report small earnings increases. For the going-concern test, the going-concern audit report variable (*GCREPORT*) is from Audit Analytics, and a total of 173 firm-year observations (2.6 percent) received going-concern reports during the sample period.

We now turn to the auditor-related control variables. *INFLUENCE* measures a client's size *relative* to the audit office size, and has a mean of 0.065, indicating that on average a client's total fees represent 6.5 percent of office-level fees. However, the median value is only 0.019 or 1.9 percent of fees. Short tenure (*TENURE*) of three years or less occurs for

---

[8] These values are larger than the means/medians reported in Table 1 for each Big 4 auditor in the full Audit Analytics population, which means that smaller offices are under-represented in the sample and larger offices are over-represented, relative to the Audit Analytics population. The most likely explanation for this is that the Compustat population is biased toward larger companies, and these companies are more likely to be audited by larger offices. The under-representation of smaller offices in the sample would bias against the predicted results by reducing the variation in the experimental variable (office size).

TABLE 2
Pooled Firm-Year Descriptive Statistics (2003–2005)

Variables	n	Mean	Median	Std. Dev.	Min.	Max.
OFFICE	6,568	65.415	40.839	84.690	0.056	623.535
lnOFFICE	6,568	17.346	17.525	1.260	10.933	20.251
ACCRUALS	6,568	0.009	0.009	0.401	−0.999	0.999
ABS_ACCRUALS	6,568	0.258	0.124	0.308	0.000	0.999
SMALL_PROFIT	6,568	0.233	0.000	NA	0.000	1.000
SMALL_INCREASE	6,568	0.129	0.000	NA	0.000	1.000
GCREPORT	6,568	0.026	0.000	NA	0.000	1.000
PRIORGC	6,568	0.026	0.000	NA	0.000	1.000
INFLUENCE	6,568	0.065	0.019	0.134	0.000	1.000
WEAKNESS	6,568	0.131	0.000	0.807	0.000	20.000
OPSEG	6,568	1.100	1.000	0.607	1.000	10.000
GEOSEG	6,568	2.415	2.000	1.930	1.000	18.000
CFO	6,568	0.064	0.113	0.345	−11.612	1.349
CFOVOLATILITY	6,568	0.264	0.098	0.825	0.004	10.000
SALESGROWTH	6,568	0.173	0.112	0.368	−1.000	2.000
SALESVOLATILITY	6,568	0.476	0.276	0.796	0.000	10.000
LOSS	6,568	0.242	0.000	NA	0.000	1.000
LAGLOSS	6,568	0.265	0.000	NA	0.000	1.000
LAGRETURN	6,568	0.895	0.324	2.055	−0.991	10.000
TENURE	6,568	0.204	0.000	NA	0.000	1.000
SIZE	6,568	6.037	6.013	1.899	−2.172	13.420
DEBT	6,568	0.483	0.456	0.487	0.002	19.963
BANKRUPTCY	6,568	2.341	2.151	3.570	−59.402	230.036
REPORTLAG	6,568	48.141	43.000	25.253	0.000	360.000
CASH	6,568	0.269	0.178	0.257	0.000	1.000
VOLATILITY	6,568	0.136	0.111	0.099	0.011	0.990
MB	6,568	1.367	0.943	2.180	−4.928	16.871
NATIONAL-LEADER	6,568	0.297	0.000	NA	0.000	1.000
CITY-LEADER	6,568	0.640	1.000	NA	0.000	1.000

Variable Definitions:

OFFICE = measure of practice office size based on aggregated client audit fees (in $ millions) of a practice office in a specific fiscal year. In the multivariate tests, log of *OFFICE* (denoted ln*OFFICE*) is used as the test variable and is based on actual fees (not rounded to millions);

ACCRUALS = signed abnormal accruals;

ABS_ACCRUALS = absolute value of abnormal accruals derived from the performance adjusted accruals model in Equation (1);

SMALL_PROFIT = dummy variable, and coded as 1 if a client's net income deflated by lagged total assets is between 0 and 5 percent, and 0 otherwise;

SMALL_INCREASE = dummy variable, and coded as 1 if a client's net income deflated by lagged total assets is between 0 and 1.3 percent, and 0 otherwise;

GCREPORT = dummy variable that takes the value of 1 if a firm receives a going-concern report in a specific fiscal year, and 0 otherwise;

PRIORGC = dummy variable that takes the value of 1 if a client received a going-concern report in the previous year, and 0 otherwise;

INFLUENCE = ratio of a specific client's total fees (audit fees plus nonaudit fees) relative to aggregate annual fees generated by the practice office which audits the client;

(*continued on next page*)

TABLE 2 (continued)

WEAKNESS = number of material internal control weaknesses reported in Audit Analytics for the firm in year *t*;

OPSEG = number of operating segments reported in the Compustat segments database for the firm in year *t*;

GEOSEG = number of geographic segments reported in the Compustat segments database for the firm in year *t*;

CFO = operating cash flows deflated by lagged total assets;

DEBT = client's total liabilities deflated by total assets;

CFOVOLATILITY = standard deviation of *CFO*. We use a rolling window and require three years of data to estimate. This variable has a winsorized maximum value of 10;

SALESGROWTH = one-year growth rate of a firm's sales revenue, and the maximum value is winsorized at 2;

SALESVOLATILITY = standard deviation of sales revenue. We use a rolling window and require three years of data. The variable has a winsorized maximum value of 10;

LOSS = dummy variable that takes the value of 1 if operating income after depreciation is negative, and 0 otherwise;

LAGLOSS = dummy variable that takes the value of 1 if operating income after depreciation in previous year is negative, and 0 otherwise;

LAGRETURN = firm's 12-month stock returns for the prior fiscal year;

TENURE = dummy variable that takes the value of 1 if auditor tenure is three years or less, and 0 otherwise;

SIZE = natural log of a client's total assets (in $ millions);

BANKRUPTCY = the Altman Z-score, which is a measure of the probability of bankruptcy, with a lower value indicating greater financial distress;

REPORTLAG = number of days between a client's fiscal year-end and its earnings announcement date;

CASH = sum of a client's total cash and investment securities deflated by total assets;

VOLATILITY = client's stock volatility and is the standard deviation of 12 monthly stock returns for the current fiscal year;

MB = natural log of the ratio of a client's market value of equity to its book value of equity;

NATIONAL-LEADER = dummy variable that takes the value of 1 if an auditor is the number one auditor in an industry in terms of aggregated audit fees in a specific fiscal year, and 0 otherwise;

CITY-LEADER = dummy variable that takes the value of 1 if an office is the number one auditor in terms of aggregated client audit fees in an industry within that city in a specific fiscal year, and 0 otherwise;

Variables are based on data in either Compustat or Audit Analytics.

20.4 percent of observations. National industry leaders (*NATIONAL-LEADER*) audit 29.7 percent of the sample, and 64 percent are audited by city-specific industry leaders (*CITY-LEADER*). Operating segment (*OPSEG*) has a mean (median) value of 1.1 (1), and geographic segment (*GEOSEG*) has a mean (median) value of 2.42 (2), though the maximum is 10 segments for *OPSEG* and 18 segments for *GEOSEG*.

Table 3 reports the correlation matrices for variables used in this study with significance levels reported at the 0.01 level. Office size is significantly correlated with "small positive earnings" in the expected direction, but is not associated with the any of the other dependent variables. Among the independent variables the primary econometric concern is the extent to which office size is correlated with the control variables. The correlation of *OFFICE* with other independent variables is quite low and less than +/− 0.15 (with one exception), which indicates that multicollinearity is not a concern in the model estimations. The exception is *OFFICE* and *INFLUENCE*, which has a Pearson (Spearman) correlation of −0.49 (−0.58). In other words, for smaller offices the client influence metric is larger, which means that the average client represents a bigger percentage of office-level fees. Given this correlation level, all models are reported both with and without the variable *INFLUENCE* to better assess if multicollinearity potentially affects the regression results, and to determine if office size is significant over and above the effect of client influence. It turns out that

TABLE 3
Correlation Matrices

Pearson Correlations above Diagonal, Spearman Correlations below Diagonal

Variables	***OFFICE***	***ACCRUALS***	***ABS_ACCRUALS***	***SMALL_PROFIT***	***SMALL_INCREASE***	***GCREPORT***	***PRIORGC***	***INFLUENCE***	***WEAKNESS***	***OPSEG***	***GEOSEG***	***CFO***	***CFOVOLATILITY***	***SALESGROWTH***
OFFICE		0.013	0.020	−0.033*	−0.022	0.004	−0.021	−0.487*	0.061*	0.010	0.120*	−0.021	0.039*	0.030
ACCRUALS	0.005		0.073*	0.046*	0.008	0.030	0.042*	0.031	−0.008	0.020	0.035*	0.021	−0.014	−0.059*
ABS_ACCRUALS	0.002	−0.067*		−0.070*	−0.084*	0.054*	0.078*	−0.059*	−0.027	−0.015	0.055*	−0.203*	0.073*	0.073*
SMALL_PROFIT	−0.046*	0.032*	−0.080*		0.140*	−0.059*	−0.034*	0.101*	0.012	0.060*	0.037*	0.049*	−0.077*	−0.077*
SMALL_INCREASE	−0.020	−0.004	−0.107*	0.140*		−0.035*	−0.039*	0.089*	−0.028	0.042*	0.005	0.068*	−0.064*	−0.048*
GCREPORT	0.006	0.031	0.045*	−0.059*	−0.035*		0.545*	−0.031	0.018	−0.012	−0.014	−0.238*	0.041*	0.009
PRIORGC	−0.022	0.041*	0.064*	−0.034*	−0.039*	0.545*		−0.066*	−0.009	−0.004	−0.035*	−0.145*	0.101*	−0.045*
INFLUENCE	−0.580*	0.020	−0.129*	0.116*	0.114*	−0.068*	−0.031*		0.033*	0.114*	0.079*	0.096*	−0.076*	−0.025
WEAKNESS	0.063*	−0.018	−0.039*	0.026	−0.022	0.010	0.011	0.078*		0.003	0.019	0.007	0.012	0.013
OPSEG	0.010	0.015	−0.018	0.036*	0.069*	−0.006	−0.001	0.143*	0.001		0.002	0.034*	−0.038*	−0.011
GEOSEG	0.136*	0.037*	0.067*	0.042*	0.017	−0.033*	−0.021	0.180*	0.020	0.019		0.071*	−0.065*	−0.040*
CFO	−0.051*	−0.010*	−0.169*	−0.083*	0.078*	−0.151*	−0.277*	0.238*	−0.027	0.060*	0.022		−0.311*	0.172*
CFOVOLATILITY	0.037	−0.020	0.265	−0.177	−0.216	0.102	0.046*	−0.325	0.026	−0.139*	−0.108*	−0.232*		0.088*
SALESGROWTH	0.029	−0.055*	0.066*	−0.096*	−0.035*	−0.040*	−0.001	−0.001	−0.002	−0.004	−0.022	−0.037*	0.117*	

(*continued on next page*)

TABLE 3 (continued)
Spearman Correlations

Variables	*OFFICE*	*ACCRUALS*	*ABS_ACCRUALS*	*SMALL_PROFIT*	*SMALL_INCREASE*	*GCREPORT*	*PRIORGC*	*INFLUENCE*	*WEAKNESS*	*OPSEG*	*GEOSEG*	*CFO*	*CFOVOLATILITY*	*SALESGROWTH*
SALESVOLATILITY	0.003	0.050*	0.064*	−0.028	−0.115*	−0.030	−0.023	−0.088*	0.040*	−0.076*	−0.028	0.008	0.441*	0.072*
LOSS	0.028	−0.049*	0.218*	−0.264*	−0.163*	0.176*	0.149*	−0.299*	0.014	−0.094*	−0.088*	−0.655*	0.428*	−0.142*
LAGLOSS	0.030	−0.050*	0.217*	−0.157*	−0.170*	0.153*	0.155*	−0.308*	−0.017	−0.070*	−0.025	−0.452*	0.457*	−0.055*
LAGRETURN	0.093*	−0.059*	−0.126*	0.026	0.073*	−0.067*	−0.036*	0.207*	0.076*	0.084*	0.046	0.085*	−0.211*	0.147*
TENURE	−0.043*	−0.004	0.013	0.011	0.001	0.009	0.006	−0.063*	−0.005	−0.016	−0.039	0.003*	0.022	−0.001
SIZE	0.111*	0.014	−0.176*	0.138*	0.196*	−0.131*	−0.155*	0.559*	0.025	0.173	0.247*	0.354*	−0.490*	0.045*
DEBT	−0.008	0.026	−0.134*	0.149*	0.104*	0.122*	0.223*	0.282*	0.030	0.105*	0.045*	0.075*	−0.288*	−0.080*
BANKRUPTCY	−0.046*	0.003	−0.074*	−0.050*	−0.031	−0.167*	−0.137*	−0.114*	−0.031	−0.050*	−0.089*	0.282*	0.010	0.075*
REPORTLAG	−0.071*	−0.037*	0.032*	0.003	−0.077*	0.173*	0.164*	−0.213*	0.129*	−0.061*	−0.245	−0.235*	0.153*	−0.046*
CASH	0.119*	−0.038*	0.266*	−0.223*	−0.174*	0.004	0.025	−0.278*	−0.015	−0.072*	0.015	−0.259*	0.493*	0.022
VOLATILITY	−0.039*	−0.020	0.216*	−0.115*	−0.231*	0.188*	0.203*	−0.319*	0.009	−0.106*	−0.071*	−0.383*	0.486*	0.005
MB	0.071*	−0.024	0.135*	−0.306*	−0.098*	0.103*	0.112*	−0.029	−0.051*	0.030	−0.001	0.158*	0.176*	0.184*
NATIONAL-LEADER	0.141*	−0.001	−0.040*	0.020	0.021	−0.009	−0.012	−0.006	−0.003	0.056*	0.067*	0.038*	−0.069*	0.005
CITY-LEADER	−0.090*	−0.012	−0.078*	0.039*	0.049*	−0.002	−0.011	0.227*	0.019	0.025	0.026	0.062*	−0.091*	−0.015

(*continued on next page*)

TABLE 3 (continued)
Pearson Correlations

Variables	*SALESVOLATILITY*	*LOSS*	*LAGLOSS*	*LAGRETURN*	*TENURE*	*SIZE*	*DEBT*	*BANKRUPTCY*	*REPORTLAG*	*CASH*	*VOLATILITY*	*MB*	*NATIONAL-LEADER*	*CITY-LEADER*
OFFICE	0.023	0.032	0.028	0.116*	−0.039*	0.133*	−0.008	−0.019	−0.056*	0.113*	−0.040*	0.035*	0.140*	−0.089*
ACCRUALS	0.048*	−0.102*	−0.093*	0.014	−0.004	0.050*	0.035*	−0.012	−0.056*	−0.115*	−0.009	−0.033*	0.001	−0.005
ABS_ACCRUALS	0.032	0.212*	0.206*	−0.062*	0.021	−0.147*	−0.002	−0.028	0.034*	0.261*	0.179*	0.098*	−0.038*	−0.055*
SMALL_PROFIT	−0.017	−0.264*	−0.158*	0.021	0.011	0.142*	0.047*	−0.006	−0.013	−0.214*	−0.112*	−0.133*	0.020	0.039*
SMALL_INCREASE	−0.062*	−0.163*	−0.169*	0.074*	0.001	0.195*	0.028	−0.009	−0.073*	−0.165*	−0.154*	−0.071*	0.021	0.049*
GCREPORT	−0.017	0.176*	0.151*	−0.026	0.009	−0.159*	0.198*	−0.133*	0.212*	0.003	0.290*	0.193*	−0.009	−0.002
PRIORGC	−0.027	0.149*	0.154*	−0.108*	0.006	−0.129*	0.099*	−0.140*	0.152*	0.023*	0.351*	0.205*	−0.012	−0.011
INFLUENCE	−0.063*	−0.160*	−0.171*	0.211*	−0.036*	0.350*	0.075*	−0.031	−0.076*	−0.165*	−0.143*	−0.013	0.017	0.160*
WEAKNESS	0.021	0.030	−0.008	0.014	0.005	0.028	0.021	−0.020	0.344*	−0.025	−0.006	0.001	0.020	0.026
OPSEG	−0.039	−0.076*	−0.051*	0.121	−0.007	0.180*	0.041*	−0.020	−0.027	−0.072*	−0.059*	0.016	0.059*	0.029
GEOSEG	−0.048*	−0.059*	−0.008	0.122	−0.028	0.220*	0.000	−0.030	−0.147*	−0.035*	−0.026	−0.008	0.059*	0.027
CFO	−0.032*	−0.502*	−0.582*	0.223*	0.020	0.382*	−0.174*	0.189*	−0.189*	−0.339*	−0.276*	−0.010*	0.045*	0.073*
CFOVOLATILITY	0.469	0.200*	0.215*	−0.026	0.016	−0.207*	−0.010	−0.031	0.042*	0.242*	0.139*	0.058*	−0.034*	−0.033*
SALESGROWTH	0.057*	−0.010	0.051*	0.013	0.011	−0.018	−0.049*	−0.046*	0.017	0.059*	0.020	0.049*	−0.005	−0.025

(*continued on next page*)

TABLE 3 (continued)

Pearson Correlations above Diagonal, Spearman Correlations below Diagonal

Variables	*SALESVOLATILITY*	*LOSS*	*LAGLOSS*	*LAGRETURN*	*TENURE*	*SIZE*	*DEBT*	*BANKRUPTCY*	*REPORTLAG*	*CASH*	*VOLATILITY*	*MB*	*NATIONAL-LEADER*	*CITY-LEADER*
SALESVOLATILITY		0.032	0.064*	−0.046*	0.026	−0.114*	−0.036*	0.057*	0.002	0.113	0.091*	−0.009	−0.009	−0.030
LOSS	0.052*		0.731*	−0.266*	−0.018	−0.461*	−0.155*	−0.101*	0.192*	0.492*	0.381*	0.092*	−0.071*	−0.089*
LAGLOSS	0.096*	0.732*		−0.138*	−0.023	−0.454*	−0.024	−0.104*	0.138*	0.507*	0.401*	0.101*	−0.075*	−0.109*
LAGRETURN	−0.123*	−0.157*	−0.241*		−0.023	0.386*	0.058*	−0.024	−0.127*	−0.086*	−0.177*	0.018	0.065*	0.086*
TENURE	−0.004	−0.017	−0.020	−0.016		−0.073*	0.021*	−0.002*	0.061*	−0.029	0.049*	0.002	−0.041*	−0.047*
SIZE	−0.208*	−0.451*	−0.466*	0.289*	−0.066*		0.061*	−0.020	−0.363*	−0.362*	−0.416*	−0.062*	0.126*	0.189*
DEBT	−0.120*	−0.014	−0.173*	0.134*	0.039*	0.372*		−0.382*	0.096*	−0.196*	0.103*	0.354*	0.025	0.034*
BANKRUPTCY	0.246*	−0.276*	−0.215*	0.029	−0.003	−0.154*	−0.540*		−0.036*	0.075*	−0.080*	−0.138*	−0.008	0.008
REPORTLAG	0.008	0.209*	0.159*	−0.120*	0.070*	−0.457*	0.028	−0.048*		−0.034*	0.221*	0.053*	−0.022	−0.042*
CASH	0.149*	0.431*	0.456*	−0.132*	−0.031	−0.342*	−0.470*	0.088*	−0.071*		0.247*	0.130*	−0.070*	−0.086*
VOLATILITY	0.259	0.451*	0.481*	−0.308*	0.050*	−0.532*	−0.132*	−0.055*	0.255*	0.320*		0.193*	−0.058*	−0.073*
MB	−0.031	0.071*	0.121*	0.076*	−0.005	−0.031	0.125*	−0.111*	−0.106*	0.298*	0.087*		−0.025	0.015
NATIONAL-LEADER	−0.030	−0.071*	−0.075*	0.035*	−0.041*	0.120*	0.069*	−0.007	−0.049*	−0.055*	−0.081*	−0.018		0.117*
CITY-LEADER	−0.023	−0.089*	−0.109*	0.069*	−0.047*	0.183*	0.123*	−0.032*	−0.075*	−0.082*	−0.099*	−0.006	0.117*	

---

* Indicates significant correlation at the 1 percent level.

All variables are defined as in Table 2.

the results on the office-size test variable are generally robust to the inclusion of *INFLUENCE* as an additional control variable.[9]

IV. RESULTS

Models are reported in this section for the accruals tests, the two benchmark earnings tests, and the going-concern reporting tests. All of the models are significant at $p < .01$, and all coefficient p-values are reported as two-tailed probabilities unless noted otherwise. Since the results are consistent in the model estimations with and without the control variable *INFLUENCE* (except as noted) we only discuss the models that include *INFLUENCE*. All sensitivity analyses in the paper are reported as two-tailed probabilities, and are based on models that control for client influence (*INFLUENCE*).

Accruals Tests

Table 4 reports the abnormal accruals test using OLS with Newey-West robust standard errors to correct for heteroscedasticity and first-order autoregressive correlation. In Panel A, for the full-sample estimation the coefficient of *lnOFFICE* is -0.011 and significant at $p < 0.01$, indicating that clients audited by larger offices have smaller abnormal accruals as hypothesized. The coefficient of *INFLUENCE* is -0.077 and significant at the 0.01 level, which is consistent with the finding of Reynolds and Francis (2000) that influential clients have smaller accruals. Five control variables are significant in the expected direction (*TENURE*, *SALESGROWTH*, *CFO*, *VOLATILITY*, *MB*), and the other control variables are not significant at the 0.10 level, or are significant in the opposite direction (*BANKRUPTCY*).

The remainder of Table 4 reports the subsample of observations having negative abnormal accruals in Panel B (note that the absolute value of negative abnormal accruals is used to create consistent predictions), and the subsample of observations with positive abnormal accruals in Panel C. For both subsamples, the coefficient of *lnOFFICE* is negative and significant at the 0.01 level, while *INFLUENCE* is not significant at the 0.10 level. In sum, the results on negative- and positive-signed accruals are consistent with the full-sample results using unsigned abnormal accruals. We conclude that firms audited by larger offices have both smaller income-increasing accruals and smaller income-decreasing accruals. Virtually identical results are obtained if a restricted maximum likelihood (RML) method is used to estimate a linear mixed model with multilevel random effects in which industry membership treated as a second-level random effect (Goldstein 1995).

We estimate the economic significance of office-size effect as follows. When *lnOFFICE* increases in size from the 25th to the 75th percentile value, unsigned abnormal accruals decrease by 0.027 (setting other model variables equal to their median values). We divide 0.027 by the sample median operating income after depreciation to estimate the increase in office size on median earnings. The magnitude is equal to a change of 8.9 percent in

---

[9] The one exception is that office size is only weakly significant at $p = .085$ (one-tailed) in the benchmark target of reporting small profits when *INFLUENCE* is included, compared to $p < .01$ when it is excluded. Given the large correlation between office size and client influence, we undertake an additional analysis by dropping the largest 25 percent and smallest 25 percent of observations for the variable *INFLUENCE*, and re-estimate the models on this reduced sample. The purpose of the analysis is to rule out that extreme values of client influence are driving the full-sample results. We find that all tests of office size are significant at the 0.05 level or less, except the going-concern test, which is significant at the 0.10 level (one-tailed). We conclude that the results are robust to extreme values of *INFLUENCE* and that client influence is not a confounding factor in the test of office size.

TABLE 4
Performance Adjusted Discretionary Accruals Test[a]

Panel A: Absolute Abnormal Accruals[b]

Independent Variables	Predicted Sign	Coefficient Estimate	p-value	Coefficient Estimate	p-value
Experimental Variable					
lnOFFICE	−	−0.011	0.000	−0.016	0.000
Control Variables					
INFLUENCE	−			−0.077	0.012
TENURE	+	0.021	0.014	0.020	0.017
NATIONAL-LEADER	−	0.004	0.607	0.005	0.523
CITY-LEADER	−	0.005	0.461	0.006	0.400
OPSEG	?	0.002	0.712	0.002	0.608
GEOSEG	?	0.003	0.165	0.003	0.142
SIZE	−	−0.001	0.570	0.001	0.706
SALESGROWTH	+	0.022	0.077	0.022	0.076
SALESVOLATILITY		0.008	0.104	0.008	0.112
CFO	−	−0.045	0.001	−0.046	0.000
CFOVOLATILITY	+	−0.002	0.729	−0.001	0.789
WEAKNESS	+	−0.004	0.207	−0.004	0.291
DEBT	+	−0.007	0.567	−0.006	0.612
LOSS	−	−0.001	0.938	0.000	0.989
BANKRUPTCY	−	0.002	0.007	0.002	0.008
VOLATILITY	+	0.247	0.000	0.246	0.000
MB	+	0.006	0.006	0.006	0.005
Constant		0.169	0.001	0.238	0.000
n			6,568		6,568
Number of Unique Firms			2,572		2,572
Adjusted R^2			0.407		0.388

Panel B: Absolute Negative Abnormal Accruals[c]

Independent Variables	Predicted Sign	Coefficient Estimate	p-value	Coefficient Estimate	p-value
Experimental Variable					
lnOFFICE	−	−0.011	0.005	−0.015	0.003
Control Variables					
INFLUENCE	−			−0.062	0.120
TENURE	+	0.030	0.014	0.029	0.015
NATIONAL-LEADER	−	0.003	0.770	0.004	0.701
CITY-LEADER	−	0.002	0.830	0.003	0.768
OPSEG	?	−0.003	0.625	−0.003	0.672
GEOSEG	?	0.002	0.567	0.002	0.554
SIZE	−	−0.001	0.868	0.001	0.752
SALESGROWTH	+	0.026	0.114	0.026	0.113
SALESVOLATILITY		0.000	0.970	0.000	0.975
CFO	−	0.013	0.644	0.011	0.676
CFOVOLATILITY	+	0.002	0.806	0.002	0.772
WEAKNESS	+	−0.001	0.842	0.000	0.945
DEBT	+	−0.035	0.019	−0.035	0.021
LOSS	−	0.058	0.000	0.059	0.000
BANKRUPTCY	−	0.001	0.185	0.001	0.200

(*continued on next page*)

TABLE 4 (continued)

Independent Variables	Predicted Sign	Coefficient Estimate	p-value	Coefficient Estimate	p-value
VOLATILITY	+	0.228	0.001	0.226	0.001
MB	+	0.011	0.002	0.011	0.002
Constant		0.223	0.002	0.300	0.311
n			3,111		3,111
Number of Unique Firms			1,962		1,962
Adjusted R^2			0.865		0.874

Panel C: Positive Abnormal Accruals[d]

Independent Variables	Predicted Sign	Coefficient Estimate	p-value	Coefficient Estimate	p-value
Experimental Variable					
lnOFFICE	–	−0.008	0.020	−0.012	0.007
Control Variables					
INFLUENCE	–			−0.061	0.167
TENURE	+	0.012	0.302	0.011	0.327
NATIONAL-LEADER	–	0.003	0.786	0.003	0.748
CITY-LEADER	–	0.005	0.646	0.005	0.616
OPSEG	?	0.001	0.918	0.001	0.850
GEOSEG	?	0.004	0.115	0.004	0.101
SIZE	–	0.001	0.869	0.003	0.479
SALESGROWTH	+	0.007	0.651	0.007	0.647
SALESVOLATILITY		0.012	0.082	0.012	0.087
CFO	–	−0.080	0.000	−0.081	0.000
CFOVOLATILITY	+	−0.003	0.613	−0.003	0.645
WEAKNESS	+	0.000	0.933	0.001	0.838
DEBT	+	0.023	0.024	0.024	0.018
LOSS	–	−0.071	0.000	−0.070	0.000
BANKRUPTCY	–	0.006	0.057	0.006	0.054
VOLATILITY	+	0.244	0.000	0.243	0.000
MB	+	0.000	0.867	0.000	0.858
Constant		0.157	0.103	0.211	0.042
n			3,457		3,457
Number of Unique Firms			2,137		2,137
Adjusted R^2			0.687		0.723

---

[a] This table reports the results of OLS estimation with Newey-West robust standard errors to correct for heteroscedasticity and first-order autocorrelation. All variables are defined as in Table 2.

[b] Panel A reports results for the total sample in which the dependent variable is the absolute value of performance adjusted abnormal accruals derived from Equation (1).

[c] Panel B reports results for the subsample of observations with negative abnormal accruals and the dependent variable is the absolute value of negative abnormal accruals derived from Equation (1).

[d] Panel C reports results for the subsample of observations with positive abnormal accruals, and the dependent variable is positive abnormal accruals derived from Equation (1).

pre-tax operating income after depreciation, which indicates the office size results are both economically and statistically significant.

A potential concern is that variables may be correlated with the error term but not included in the model. In order to address this classic omitted-variables threat, a firm-level fixed-effect model is estimated.[10] This is an especially important test because it provides an additional control for possible omitted client risk variables that may drive the results rather than the auditor's office size. In this estimation the coefficient of *lnOFFICE* is significant at the 0.01 level in all of the accruals tests, and provides additional assurance that omitted client risk factors do not affect the tests.

In sum, the tests consistently show that client abnormal accruals vary across offices and support the conjecture that larger offices provide higher quality audits by constraining the client's use of accruals to manage reported earnings. The economic magnitudes are large and the results are robust to alternative estimation approaches including firm fixed-effect models as an additional control for client risk factors.

Earnings Benchmark Tests

Table 5 reports the two earnings benchmark tests. We also estimate random effect probit models as an additional control for omitted firm variables, and these results are comparable to those reported in Table 5. Panel A of Table 5 reports the tests of the benchmark target of small positive earnings. The first model excludes *INFLUENCE* and the coefficient for *lnOFFICE* is -0.063 and significant at $p < .01$, indicating that clients audited by larger offices are less likely to have managed earnings in order to report small earnings (and avoid losses). Seven control variables are significant in the expected direction (*OPSEG, GEOSEG, SALESGROWTH, SALESVOLATILITY, CFO, DEBT, LOSS*), one control variable is significant in the opposite direction (*MB*), and the remaining control variables are insignificant at the 10 percent level. The second model includes *INFLUENCE*, and the coefficients of both *lnOFFICE* and *INFLUENCE* are insignificant at the 10 percent level. However, *lnOFFICE* is weakly significant in a one-tailed test ($p = .085$). In addition, office size is significant at the 0.01 level (with *INFLUENCE* in the model), when measured as the rank of office size in which audit fees are used to rank offices. In sum, there is evidence that clients in larger offices are less likely to meet the earnings benchmark target of reporting small positive earnings, although the results are weaker when the client influence variable (*INFLUENCE*) is included.

Panel B of Table 5 reports the tests of small earnings increases. The coefficient on *lnOFFICE* is -0.06, and is significant at $p = .013$, while *INFLUENCE* is insignificant at the 0.10 level. Two control variables are significant in the expected direction (*CFO, LOSS*), five control variables are significant in the opposite direction (*SIZE, SALESGROWTH, WEAKNESS, VOLATILITY, MB*), and the remaining control variables are insignificant at the 10 percent level. We conclude that clients in larger offices are less likely to meet the benchmark target of reporting small earnings increases.

Following Denis and Mihov (2003), we estimate the hypothetical changes in the probability of meeting benchmarks when *lnOFFICE* changes from its 25th percentile to 75th percentile value, while the other independent variables remain constant at their mean values. For the small positive earnings test, the increase in office size decreases the probability by 1.0 percent (from 11.30 percent at the 25th percentile to 10.30 percent at the 75th percentile). For the small earnings increase test, the probability decreases by 1.39 percent (from 8.73 percent to 7.34 percent). These amounts are modest in absolute magnitudes,

---

[10] Greene (2007, Chap. 14) indicates that both fixed-effect and random-effect models control for omitted variables in panel data, and that a fixed-effect model is a special case of the more general random-effect model. For the accruals test in our sample, a Hausman specification test shows that a fixed-effect model is more appropriate than a random-effect model, although we get virtually the same results if we use a random-effect model. For the earnings benchmark tests and going-concern report tests, we only report robustness to estimations using random-effect models as there is not a fixed-effect probit model (Wooldridge 2002).

TABLE 5
Earnings Benchmark Tests[a]

Panel A: Reporting Small Positive Earnings[b]

Independent Variables	Predicted Sign	Coefficient Estimate	p-value	Coefficient Estimate	p-value
Experimental Variable					
lnOFFICE	−	−0.063	0.006	−0.039	0.171
Control Variables					
INFLUENCE	−			0.354	0.166
TENURE	+	0.053	0.411	0.057	0.375
NATIONAL-LEADER	−	−0.042	0.502	−0.045	0.472
CITY-LEADER	−	0.040	0.497	0.039	0.515
OPSEG	?	0.112	0.009	0.109	0.011
GEOSEG	?	0.030	0.044	0.029	0.056
SIZE	−	0.021	0.275	0.009	0.661
SALESGROWTH	+	−0.800	0.000	−0.798	0.000
SALESVOLATILITY	+	0.120	0.010	0.119	0.010
CFO	−	−0.664	0.000	−0.660	0.000
CFOVOLATILITY	+	−0.231	0.001	−0.233	0.001
WEAKNESS	+	0.048	0.122	0.045	0.150
DEBT	+	0.095	0.186	0.092	0.198
LOSS	−	−2.021	0.000	−2.025	0.000
BANKRUPTCY	−	−0.029	0.151	−0.029	0.149
VOLATILITY	+	−0.007	0.983	0.010	0.975
MB	+	−0.148	0.000	−0.148	0.000
Constant		0.402	0.335	0.048	0.922
n			6,568		6,568
Number of Unique Firms			2,572		2,572
Pseudo R^2			0.118		0.119
Percent Concordant			0.722		0.736
Percent Discordant			0.278		0.264

Panel B: Reporting Small Increases in Earnings[c]

Independent Variables	Predicted Sign	Coefficient Estimate	p-value	Coefficient Estimate	p-value
Experimental Variable					
lnOFFICE	−	−0.048	0.013	−0.060	0.013
Control Variables					
INFLUENCE	−			−0.175	0.405
TENURE	+	0.051	0.372	0.049	0.393
NATIONAL-LEADER	−	−0.004	0.945	−0.002	0.972
CITY-LEADER	−	0.032	0.531	0.034	0.515
OPSEG	?	0.034	0.307	0.036	0.283
GEOSEG	?	−0.018	0.164	−0.017	0.188
SIZE	−	0.110	0.000	0.116	0.000
SALESGROWTH	+	−0.333	0.000	−0.334	0.000
SALESVOLATILITY	+	−0.058	0.158	−0.058	0.157
CFO	−	−0.179	0.043	−0.182	0.040
CFOVOLATILITY	+	−0.013	0.788	−0.012	0.811
WEAKNESS	+	−0.071	0.070	−0.069	0.078
DEBT	+	0.083	0.245	0.085	0.236

(*continued on next page*)

TABLE 5 (continued)

Independent Variables	Predicted Sign	Coefficient Estimate	p-value	Coefficient Estimate	p-value
LOSS	–	−0.524	0.000	−0.521	0.000
BANKRUPTCY	–	−0.015	0.420	−0.014	0.423
VOLATILITY	+	−2.050	0.000	−2.055	0.000
MB	+	−0.070	0.000	−0.070	0.000
Constant		−0.654	0.061	−0.475	0.246
n			6,568		6,568
Number of Unique Firms			2,572		2,572
Pseudo R^2			0.107		0.108
Percent Concordant			0.713		0.717
Percent Discordant			0.287		0.283

---

[a] This table reports results of probit model estimates of Equation (3) with clustered robust standard errors to correct for heteroscedasticity and serial dependence. All variables are defined as in Table 2.

[b] Panel A reports results of reporting small positive earnings. We code an observation as 1 if it reports small positive earnings (net income deflated by lagged total assets is between 0 and 5 percent), and 0 otherwise.

[c] Panel B reports results of reporting small increases in earnings. We code an observation as 1 if it reports small earnings increases (change in net income deflated by lagged total assets is between 0 and 1.3 percent), and 0 otherwise.

although in percentage terms they represent an 8.8 percent reduction in the likelihood of reporting small positive earnings, and a 15.9 percent reduction in the likelihood of reporting small positive earnings changes.

In summary, results in this section show that clients audited by larger offices are less likely to report small earnings and small earnings increases that are consistent with less earnings management behavior. Results are robust to alternative estimations using random-effect models as an additional control for omitted client variables, and the economic magnitudes are relatively large in terms of percentage changes.

Going Concern Tests

Table 6 reports the going-concern tests. A probit model is estimated with robust standard errors to correct for heteroscedasticity and serial dependence. We also estimate a random-effect probit model as an additional control for omitted client variables, and the results are consistent with those reported in Table 6. The models in Table 6 are consistent both with and without the control variable *INFLUENCE*, so we only discuss the results including *INFLUENCE*. Panel A reports tests using the full sample. The coefficient on *lnOFFICE* is 0.112 and is significant at p = .013, while the coefficient of *INFLUENCE* is insignificant at the 0.10 level. Nine control variables are significant in the expected direction (*CITY-LEADER*, *SIZE*, *CASH*, *PRIORGC*, *REPORTLAG*, *LOSS*, *BANKRUPTCY*, *VOLATILITY*, *MB*), and the other control variables are insignificant at the 10 percent level except *DEBT* and *LAGRETURN*, which are significant in the opposite than expected direction. These results indicate that auditors in larger offices are more likely to issue going-concern reports, *ceteris paribus*.

Some studies limit going-concern analysis to subsamples of financially distressed clients arguing that a going-concern opinion decision is most salient for financially distressed clients (Hopwood et al. 1994; Mutchler et al. 1997; Reynolds and Francis 2000; DeFond et al. 2002). Therefore, as an additional test the sample is limited to 2,022 financially

TABLE 6
Going Concern Audit Report Tests[a]

Panel A: Total Sample[b]

Independent Variables	Predicted Sign	Coefficient Estimate	p-value	Coefficient Estimate	p-value
Experimental Variable					
lnOFFICE	+	0.089	0.024	0.112	0.013
Control Variables					
INFLUENCE	+			0.583	0.269
TENURE	+	0.013	0.910	0.015	0.890
NATIONAL-LEADER	+	−0.049	0.654	−0.054	0.616
CITY-LEADER	+	0.228	0.026	0.229	0.026
OPSEG	?	−0.052	0.628	−0.052	0.622
GEOSEG	?	0.012	0.618	0.011	0.650
SALESVOLATILITY	+	−0.034	0.667	−0.036	0.651
SIZE	−	−0.111	0.000	−0.124	0.000
CASH	−	−1.259	0.000	−1.271	0.000
PRIORGC	+	1.742	0.000	1.743	0.000
REPORTLAG	+	0.006	0.000	0.006	0.000
DEBT	+	−0.118	0.063	−0.116	0.067
LOSS	+	0.617	0.000	0.615	0.000
LAGLOSS	+	0.135	0.342	0.139	0.329
BANKRUPTCY	−	−0.043	0.038	−0.041	0.048
LAGRETURN	−	0.049	0.036	0.047	0.044
VOLATILITY	+	1.211	0.000	1.222	0.000
MB	+	0.065	0.000	0.064	0.000
Constant		−3.855	0.000	−4.202	0.000
n			6,568		6,568
Number of Unique Firms			2,572		2,572
Pseudo R^2			0.475		0.471
Percent Concordant			0.891		0.898
Percent Discordant			0.109		0.102

Panel B: Financially Distressed Clients[c]

Independent Variables	Predicted Sign	Coefficient Estimate	p-value	Coefficient Estimate	p-value
Experimental Variable					
lnOFFICE	+	0.085	0.056	0.101	0.045
Control Variables					
INFLUENCE	+			0.468	0.491
TENURE	+	−0.021	0.869	−0.020	0.877
NATIONAL-LEADER	+	0.004	0.973	0.001	0.995
CITY-LEADER	+	0.297	0.009	0.297	0.009
OPSEG	?	−0.011	0.933	−0.011	0.930
GEOSEG	?	−0.001	0.978	−0.001	0.959
SALESVOLATILITY	+	−0.078	0.390	−0.080	0.381
SIZE	−	−0.149	0.000	−0.158	0.000
CASH	−	−1.301	0.000	−1.309	0.000
PRIORGC	+	1.770	0.000	1.769	0.000
REPORTLAG	+	0.005	0.000	0.005	0.000

(*continued on next page*)

TABLE 6 (continued)

Independent Variables	Predicted Sign	Coefficient Estimate	p-value	Coefficient Estimate	p-value
DEBT	+	−0.103	0.113	−0.102	0.118
LOSS	+	−0.164	0.319	−0.164	0.317
LAGLOSS	+	0.072	0.644	0.075	0.629
BANKRUPTCY	−	−0.031	0.151	−0.030	0.170
LAGRETURN	−	0.099	0.001	0.097	0.001
VOLATILITY	+	1.732	0.000	1.737	0.000
MB	+	0.040	0.009	0.039	0.011
Constant		−2.825	0.000	−3.069	0.000
n			2,022		2,022
Number of Unique Firms			1,145		1,145
Pseudo R^2			0.389		0.399
Percent Concordant			0.885		0.887
Percent Discordant			0.115		0.113

---

[a] This table reports results of probit model estimates of Equation (4) with clustered robust standard errors to correct for heteroscedasticity and serial dependence. *GCREPORT* is the independent variable in all tests. All variables are defined as in Table 2.

[b] Panel A reports results based on the total sample.

[c] Panel B reports results for a subsample of financially distressed clients (firms with non-positive net income).

distressed firm-year observations (firms with non-positive net income) and these results are reported in Panel B of Table 6. The coefficient of *lnOFFICE* is 0.101 and significant at p = .045, while *INFLUENCE* is insignificant at the 0.10 level. Thus, the results on *lnOFFICE* are consistent in both the full sample and the distressed subsample, and indicate that auditors in larger offices are more likely to issue going-concern reports.

We use the same approach as in the previous section to measure changes in the probability of issuing going-concern reports for hypothetical changes in office size. When we use the full sample, the increase in office size from the 25th to 75th percentile value increases the probability of issuing a going-concern report by 0.28 percent (from 0.40 percent at the 25th percentile to 0.68 percent at the 75th percentile). In the distressed subsample it increases the probability by 1.26 percent (from 2.85 to 4.11). While these amounts seem small in absolute magnitudes, the percentage effects are quite large. For example, the increase of 1.26 percent in the distressed subsample represents a 44 percent increase in the going concern rate.

The final analysis examines the accuracy of the audit opinion. The finding that larger offices are more likely to issue a going-concern report may simply reflect a more conservative reporting policy by larger offices rather than greater expertise. To assess audit report accuracy, we identify those firms in the sample that declared bankruptcy in the year following the audit report by examining footnotes 33, 34, and 35 in Compustat. In order to conduct Chi-square tests, we classify offices into "small" and "large" based on the median value of office-level fees (*OFFICE*). A total of 173 going-concern reports were issued in the sample for which nine clients went bankrupt the next year. In terms of reporting accuracy, "large" offices were more accurate in predicting client failure. Large offices issued 90 going-concern reports and had eight client failures, while small offices issued 83 going-concern reports and had just one client failure. A Chi-square test is significant at the 5 percent level and indicates that larger offices are better at identifying those clients likely to fail in the next year, although both auditor groups over-issue going-concern reports relative to actual next-period bankruptcies.

In summary, results in this section show that larger offices are more likely to issue going-concern reports, and these reports are more accurate in predicting next-period bankruptcy. These results are consistent with those documented in the abnormal accruals tests and earnings benchmark tests. Together the results across all tests support the argument that there is significant variation in audit outcomes across different sizes of Big 4 practice offices, with the evidence consistently supporting that larger offices provide higher quality audits.

V. SENSITIVITY AND ROBUSTNESS TESTS

Control for Client Risk Factors

A potential concern is that office size may capture omitted client risk factors, and that it is these omitted client risk factors that drive the results rather than size of the Big 4 audit office. We cannot entirely rule this out but the models in Tables 4–6 include a comprehensive set of variables that control for multiple dimensions of client risk. Operating risk is controlled by variables such as sales volatility and sales growth, cash flow from operation and volatility of cash flows, the number of operating and geographic segments, and internal control deficiencies reported under SOX. Financing risk is controlled with variables for client size, profitability (reporting a loss in the current and prior period), leverage, and bankruptcy probability. Market risk is controlled with variables for market-to-book ratio, stock returns, and the standard deviation of stock returns. In addition, as reported in Section IV, office size is significant in all tests at comparable levels when fixed- and random-effect models are estimated as an additional control for potential omitted client risk factors.

To determine the extent to which the study's client risk factors are systematically related to auditor office size, we regress *lnOFFICE* on all 15 of the above client risk variables along with an additional control for industry fixed effects. The adjusted R^2 in this estimation is only 7.3 percent, indicating that auditor office size is largely unexplained by client-specific and client-industry risk factors. This can also be seen in Table 3 as only five variables have a Pearson correlation with office size that is 0.10 or greater: *GEOSEG*, *LAGRETURN*, *SIZE*, *CASH*, and *NATIONAL-LEADER*. As an additional sensitivity, an interaction term is created for *lnOFFICE* and each of these five variables, and added to the models in Tables 4–6. None of the interaction terms is significant at the 0.10 level, and the test variable *lnOFFICE* remains significant at levels reported in the tables, which provides additional evidence the results are not driven by these particular client risk variables.

Other Sensitivity Analyses

To assure the results on *lnOFFICE* are not driven by extreme values of office size, we drop the bottom and top 10 percent of firm-year observations having the most extreme values of office size (based on audit fees). For this trimmed sample, the results for *lnOFFICE* are significant at the 0.05 level or less across all tests in Tables 4–6.

An alternative office-size metric is tested using the natural log of an office's total fees (the sum of audit and nonaudit fees). These untabulated results are consistent with those reported in Tables 4 through 6 based on log of audit fees. We also use the log of office ranks as an alternative measure of office size, where the 805 offices in the sample are ranked on their office-level audit fees. Results on this measure of office size are significant in the expected direction at the 0.05 level or less across all tests.

Another concern is that smaller clients may be more likely to be audited by smaller offices. If this is the case and if smaller clients have inherently lower quality earnings, then client size (rather than office size) may drive the results. To rule this out, Table 7 reports results based on separate subsample estimations of all models for "large" and "small" clients. Construction of the large/small partition splits each year's observations into upper/lower halves based on yearly median values of client assets.

In Table 7, office size is significant in the small and large subsamples in all models estimations of the accruals tests ($p < .01$) and the benchmark tests of reporting small earnings increases ($p < .08$), indicating that client size does not drive the results on these tests. In the other two tests, there are differences between the two subsamples. For the benchmark test of reporting small positive earnings, the coefficient of *lnOFFICE* is significant at $p < .10$ in all model estimations for large clients, but is insignificant for small clients indicating the results in Table 5 are not generalizable to the full sample but instead are limited to the *larger* half of sample observations. An explanation for this result is that large firms have stronger incentives to meet earnings benchmarks due to more widespread shareholdings (and are more constrained by large offices), whereas smaller firms may be more closely held and less driven by short-term market forces to meet such benchmarks. For the going-concern tests, the coefficient of *lnOFFICE* is significant at $p < .05$ for small clients, but is *insignificant* for large clients. In large part this result is due to the fact that nearly all of the going-concern reports (134 of the 173) are in the "small" client segment, which lowers the power of the test for "large" clients.

In sum, the results in Table 7 indicate two of the four tests are consistent for the subsamples of small and large firms. For the two tests that differ, there is no evidence of a systematic client-size effect since office size is significant in the going-concern model only for *smaller* clients, whereas office size is significant in the small profit test only for *larger* clients.

Auditor Effects

This section discusses the auditor-related control variable *TENURE* and indicator variables for national-level and city-specific industry leadership. These variables are intended to control for other auditor effects that could be correlated omitted variables.[11]

The short tenure variable (*TENURE*) is positively associated with unsigned accruals as in Johnson et al. (2002), and with negative abnormal accruals. However, *TENURE* is insignificant in all other tests.

Regarding industry expertise, city-specific industry leadership is positively associated with going-concern audit reports, which is consistent with greater expertise by industry experts. Otherwise, the national- and city-level industry leadership variables are insignificant. These results do not appear to support the arguments and findings in prior research that national industry leadership (Balsam et al. 2003) or city-specific industry leadership (Francis et al. 2006) affect the quality of reported earnings by clients. However, these prior studies are based on samples from earlier periods, and do not include variables for office

---

[11] In addition, to rule out that office size captures accounting firm effects (a possible omitted variable), all models in Tables 4–6 are re-estimated with indicator variables for each of the Big 4 accounting firms (technically three indictor variables, with the fourth audit firm captured by default in the intercept). None of the Big 4 variables is significant in any of the tests, indicating no auditor-specific associations with the dependent variables. In addition, the test variable *lnOFFICE* remains significant at comparable levels when the accounting firm dummies are added with one exception: office size is not significant at the 0.10 level in the test of positive abnormal accruals.

TABLE 7
Sensitivity Analyses Based on Upper and Lower Halves of Yearly Client Size

	Large Clients		Small Clients	
	Coefficient Estimate	p-value	Coefficient Estimate	p-value
Accruals Tests				
OLS Estimates with Robust Standard Errors and Industry Dummies	−0.019	0.000	−0.014	0.005
RML Estimates with Multilevel Random Effects	−0.019	0.000	−0.017	0.001
Firm-Level Fixed-Effect Model	−0.103	0.000	−0.089	0.000
Earnings Benchmark Tests				
Reporting Small Profits				
Probit with Robust Standard Errors	−0.064	0.046	−0.009	0.775
Random-Effect Probit Model	−0.070	0.089	−0.020	0.604
Reporting Small Earnings Increases				
Probit with Robust Standard Errors	−0.073	0.022	−0.059	0.074
Random-Effect Probit Model	−0.078	0.022	−0.059	0.080
Going-Concern Tests				
Total Sample	−0.063	0.565	0.131	0.016
Financially Distressed Firms	−0.188	0.542	0.122	0.038

This table reports separate estimation results for large clients and small clients in which the total sample is split based on the yearly median values of client size (assets). Note that the variable for client influence (*INFLUENCE*) is included in the estimations summarized in this table, along with all other control variables in Tables 4–6.

size or client influence. Therefore we re-estimate all models by dropping the variables *lnOFFICE* and *INFLUENCE*. These new estimations show that city-level industry leadership is negative and significant at the 5 percent level in all of the accruals tests, which is consistent with Francis et al. (2006), and also shows a positive association with going-concern reports (consistent with Table 6). However, city leadership is insignificant in both of the earnings benchmark tests, and the national leadership variable is not significant in any of the tests at the 0.10 level. This analysis suggests that office size may be more important than auditor industry leadership in explaining earnings quality, but further investigation is needed to tease out and better understand the inter-relation between office size and industry expertise.

Client Count as an Alternative Office-Size Measure

The primary tests use fee-based measures of office size because we believe fees are a good proxy for audit effort and the overall level of an office's SEC engagement experience. However, a potential problem with the fee-based approach is that fees are influenced by client characteristics such size, risk, and complexity, and these factors could be related to the study's dependent variables but may be insufficiently controlled for in the models. While the models have an extensive set of client controls, we cannot rule out that the fee-based office-size effect we document is the consequence of omitted client risk factors that are confounded with audit fees.

An alternative approach is to measure office size by the number of its SEC clients. Client count data has the potential advantage of measuring office size more independently of client characteristics that may confound the fee-based tests.[12] The Pearson (Spearman) correlation of fees and client count for the 805 offices is 0.75 (0.85) indicating that the two measures produce a similar but not identical ordering of Big 4 office size.

The models in Tables 4–6 are re-estimated using two client count measures of office size: log of clients, and log of ranks of clients (with offices ranked from one to 805 based on their client count). The accruals tests in Table 4 are significant at $p < .05$ when re-estimated with ranks of offices, but are insignificant at the 0.10 level using log of clients. Therefore, the accruals tests are robust to the measurement of office size with client-count data but are more sensitive to functional form than is the case with fee-based measures. However, the other three tests reported in Tables 5 and 6 (earnings benchmark tests and going-concern tests) are not significant at the 0.10 level in either specification, and therefore are not robust to using client-count data to measure office size.

A potential problem with the client-count approach is the large number of "ties" in client-count data. For the full sample, 733 of the 805 offices have ties in their client-count data with one or more offices, the largest of which are 46 offices with 5 clients, 38 offices with a single client, and 34 offices with 7 clients. This means that each office with the same client count is treated as identical in size and this may lead to a noisy ordering of office expertise when clients in these offices require substantially different audit effort. To illustrate this consider the 38 offices in the study that have a single SEC registrant. The SEC client in the largest office paid an audit fee of $11.4 million, which is 203.5 times larger than the smallest office's fee of $56,000: it seems unlikely auditors in these two offices would develop comparable levels of SEC engagement experience.

We control for the potential noise due to ties by re-estimating the models for 72 offices (1,832 firm-year observations) with no ties in client count. For this reduced sample, and using log of ranks of client count to measure office size, two additional tests are significant at the 0.05 level: the benchmark test of reporting small profits and the going-concern test. We conclude that the study's full-sample results are robust to using client-count data to measure office size only for the abnormal accruals tests, but when we control for potential noise by eliminating those offices with ties in client count two additional tests are significant.

VI. CONCLUSION

We document a systematic association between Big 4 office size and audit outcomes consistent with larger offices producing higher quality audits. Results are robust to multiple specifications of office size using office-level fees. However, the results are less robust to using client count to measure office size and are sensitive to the large number of ties in office-level client-count data. While the overall evidence in the study is suggestive of an office size/audit quality relation, as a caveat we recognize that we cannot conclusively separate office-size effects from client characteristics, despite the large number of controls in the models.

Corporate scandals such as WorldCom, the failure of Arthur Andersen, and passage of the Sarbanes-Oxley Act of 2002 have created incentives for accounting firms to re-evaluate their procedures to ensure that the quality of audit work is of a high and consistent standard.

---

[12] However, this may not be the case, at least for the risk variables in our study. When we regress log of clients on the study's 15 client risk control variables (plus industry fixed effects), analogous to the regression we report above using log of office fees as the dependent variable, the R^2 is 7.1 percent using log of clients as the dependent variable, which is similar to the R^2 of 7.3 percent reported above using log of office fees. Thus, an office's client count and audit fees have similar associations with these particular client risk characteristics.

Consistent audit quality is a particular concern given that PCAOB inspection reports identified numerous deficiencies on individual audit engagements for each of the Big 4 accounting firms in their initial inspections.[13]

If local offices now engage in more consultation with the national office, have better-aligned incentives in partner compensation contracts, and enforce stricter and consistent audit procedures, then it is reasonable to expect that variation in audit quality across practice offices would have been eliminated in the post-SOX environment. Given that this does not appear to be the case, future research could usefully investigate the frictions that cause this, and the additional factors needed to mitigate the cross-sectional variation we observe in audit outcomes across practice offices of different sizes.

REFERENCES

Altman, E. 1983. *Corporate Financial Distress: A Complete Guide to Predicting, Avoiding, and Dealing with Bankruptcy*. New York, NY: Wiley.

Ashbaugh, H., R. LaFond, and B. Mayhew. 2003. Do nonaudit services compromise auditor independence? Further evidence. *The Accounting Review* 78: 611–639.

Balsam, S., J. Krishnan, and J. S. Yang. 2003. Auditor industry specialization and earnings quality. *Auditing: A Journal of Practice & Theory* 22: 71–97.

Becker, C., M. DeFond, J. Jiambalvo, and K. R. Subramanyam. 1998. The effect of audit quality on earnings management. *Contemporary Accounting Research* 15: 1–24.

Becker, G. S. 1993. *Human Capital*. 3rd edition. Chicago, IL: University of Chicago Press.

Burgstahler, D., and I. Dichev. 1997. Earnings management to avoid earnings decreases and losses. *Journal of Accounting and Economics* 24: 99–126.

Carcello, J., R. Hermanson, and N. McGrath. 1992. Audit quality attributes: The perceptions of audit partners, preparers, and financial statement users. *Auditing: A Journal of Practice & Theory* 11: 1–15.

———, D. Hermanson, and F. Huss. 1995. Temporal changes in bankruptcy-related reporting. *Auditing: A Journal of Practice & Theory* 14: 133–143.

Carey, P., and R. Simnett. 2006. Audit partner tenure and audit quality. *The Accounting Review* 81: 653–676.

Choi, J., F. Kim, J. Kim, and Y. Zang. 2007. Audit office size, audit quality and audit pricing. Working paper, Seoul National University, City University of Hong Kong, Hong Kong Polytechnic University, and Singapore Management University.

Craswell, A. T., D. J. Stokes, and J. Laughton. 2002. Auditor independence and fee dependence. *Journal of Accounting and Economics* 33: 253–275.

Danos, P., J. W. Eichenseher, and D. L. Holt. 1989. Specialized knowledge and its communication in auditing. *Contemporary Accounting Research* 6: 91–109.

DeAngelo, L. 1981. Auditor size and audit quality. *Journal of Accounting and Economics* 3: 183–199.

Dechow, P. M., R. Sloan, and A. Sweeney. 1995. Detecting earnings management. *The Accounting Review* 70: 193–225.

DeFond, M., and J. Jiambalvo. 1994. Debt covenant violation and manipulation of accruals. *Journal of Accounting and Economics* 17: 145–176.

———, K. Raghunandan, and K. R. Subramanyam. 2002. Do non-audit service fees impair auditor independence? Evidence from going concern audit opinions. *Journal of Accounting Research* 40: 1247–1274.

Degeorge, F., J. Patel, and R. Zeckhauser. 1999. Earnings management to exceed thresholds. *Journal of Business* 72: 1–33.

---

[13] Reports are available at http://www.pcaob.org/inspections/public_reports/index.aspx.

Denis, D. J., and V. T. Mihov. 2003. The choice among bank debt, non-bank private debt, and public debt: Evidence from new corporate borrowings. *Journal of Financial Economics* 70: 3–28.

Doyle, J., G. Weili, and S. McVay. 2007. Accruals quality and internal control over financial reporting. *The Accounting Review* 82: 1141–1170.

Ferguson, A., J. Francis, and D. Stokes. 2003. The effects of firm-wide and office-level industry expertise on audit pricing. *The Accounting Review* 78: 429–448.

Francis, J. R., E. Maydew, and H. C. Sparks. 1999. The role of Big 6 auditors in the credible reporting of accruals. *Auditing: A Journal of Practice & Theory* 18: 17–34.

———, D. J. Stokes, and D. J. Anderson. 1999. City markets as a unit of analysis in audit research and the re-examination of Big 6 market shares. *Abacus* 35: 185–206.

———, K. R. Reichelt, and D. Wang. 2005. The pricing of national and city-specific reputations for industry expertise in the U.S. audit market. *The Accounting Review* 80: 113–136.

———, ———, and ———. 2006. National versus office-specific measures of auditor industry expertise and effects on client earnings quality. Working paper, University of Missouri, Louisiana State University, and University of Nebraska.

Frankel, R. M., M. F. Johnson, and K. K. Nelson. 2002. The relation between auditors' fees for nonaudit services and earnings management. *The Accounting Review* 77: 71–105.

Goldstein, H. 1995. *Multilevel Statistical Models.* London, U.K.: Edward Arnold; New York, NY: Halstead Press.

Greene, W. 2007. *Econometric Analysis.* 6th edition. Upper Saddle River, NJ: Prentice Hall.

Hopwood, W., J. McKeown, and J. Mutchler. 1994. Reexamination of auditor versus model accuracy within the context of the going-concern opinion decision. *Contemporary Accounting Research* 11: 295–310.

Hribar, P., and D. Nichols. 2007. The use of unsigned earnings quality measures in tests of earnings management. *Journal of Accounting Research* 44: 1017–1053.

Johnson, V., I. Khurana, and J. K. Reynolds. 2002. Audit firm tenure and the quality of accounting earnings. *Contemporary Accounting Research* 19: 637–660.

Jones, J. 1991. Earnings management during import relief investigations. *Journal of Accounting Research* 29: 193–228.

Kothari, S. P., A. J. Leone, and C. E. Wasley. 2005. Performance matched discretionary accrual measures. *Journal of Accounting and Economics* 39: 163–197.

Levitt, A. 1998. The numbers game. Speech delivered at New York University, Center for Law and Business, September 28. Available at: http://www.sec.gov/news/speech/speecharchive/1998/spch220.txt.

Matsumoto, D. A. 2002. Management's incentives to avoid negative earnings surprises. *The Accounting Review* 77: 483–514.

Menon, K., and D. Williams. 2004. Former audit partners and abnormal accruals. *The Accounting Review* 79: 1095–1118.

Mutchler, J., W. Hopwood, and J. McKeown. 1997. The influence of contrary information and mitigating factors in audit opinion decisions on bankrupt companies. *Journal of Accounting Research* 35: 295–310.

Newey, W. K., and K. West. 1987. A simple, positive semi-definite heteroscedasticity and autocorrelation consistent covariance matrix. *Econometrica* 55: 703–708.

Raghunandan, K., and D. Rama. 1995. Audit opinions for companies in financial distress: Before and after SAS No. 59. *Auditing: A Journal of Practice & Theory* 14: 50–63.

Reynolds, J. K., and J. R. Francis. 2000. Does size matter? The influence of large clients on office-level auditor reporting decisions. *Journal of Accounting and Economics* 30: 375–400.

Rogers, W. H. 1993. Regression standard errors in clustered samples. *Stata Technical Bulletin* 13 (May): 19–23.

Satava, D. 2003. The A to Z of keeping staff. *Journal of Accountancy* 195: 67–70.

Vera-Muñoz, S., J. Ho, and C. Chow. 2006. Enhancing knowledge sharing in public accounting firms. *Accounting Horizons* 20: 133–155.

Wallman, S. 1996. The future of accounting, Part III: Reliability and auditor independence. *Accounting Horizons* 10: 76–97.
Warfield, T., J. Wild, and K. Wild. 1995. Managerial ownership, accounting choices, and informativeness of earnings. *Journal of Accounting and Economics* 20: 61–91.
Wooldridge, J. M. 2002. *Econometric Analysis of Cross-Section and Panel Data.* Cambridge, MA: MIT Press.

THE ACCOUNTING REVIEW
Vol. 84, No. 5
2009
pp. 1553–1573

American Accounting Association
DOI: 10.2308/accr.2009.84.5.1553

Accounting Discretion, Horizon Problem, and CEO Retirement Benefits

Paul Kalyta
McGill University

ABSTRACT: Empirical research on the impact of managerial retirement on discretionary accounting choices is inconclusive, with most studies finding no evidence of earnings management in the pre-retirement period. I argue that income-increasing accounting choices in final pre-retirement years are particularly appealing to managers whose pension depends on firm performance in these years. Using primary data on retired CEOs of *Fortune* 1000 firms, I investigate the impact of CEO pension plans on discretionary accruals. Consistent with the prediction, I find evidence of income-increasing earnings management in the pre-retirement period *only* when CEO pension is based on firm performance. I also report evidence of negative abnormal market reaction to CEO retirement in firms with performance-contingent CEO pensions.

Keywords: *accounting discretion; earnings management; horizon problem; executive compensation; retirement; pensions.*

Data Availability: *All primary and secondary data used in the study are publicly available.*

I. INTRODUCTION

The literature on the impact of voluntary managerial departures on discretionary accounting choices is inconclusive. In theory, the manager who plans to leave the firm has fewer incentives to act in the best interest of that firm than the manager who does not plan to leave—a phenomenon commonly referred to as the *horizon problem* (e.g., Smith and Watts 1982; Dechow and Sloan 1991; Cheng 2004). The horizon problem can affect the firm in numerous ways (including, but not limited to, earnings management, suboptimal investments, and accounting fraud) and should be especially pronounced if the manager plans to retire rather than leave the firm and stay on the job market (Gibbons and Murphy 1992). However, empirical evidence on the association between managerial retirement and discretionary choices is mixed, with most studies finding little or no support of

---

I thank Julie Marcel and Kim Nguyen for excellent assistance with data collection and analysis. I am indebted to Michel Magnan for his valuable contribution to the study. I am grateful to David Balkin, Emilio Boulianne, Dan Dhaliwal, Steve Fortin, José Garrido, Steven Kachelmeier, and anonymous reviewers for their insightful suggestions. I acknowledge financial support from *Fonds québécois de la recherche sur la société et la culture* and the Telfer School of Management, University of Ottawa. The study has benefited from comments of discussants and participants at general and regional conferences of the American Accounting Association and the Administrative Sciences Association of Canada (best paper).

Editor's note: Accepted by Steven Kachelmeier, with thanks to Dan Dhaliwal for serving as editor on a previous version.

Submitted: December 2007
Accepted: December 2008
Published Online: September 2009

the theoretical predictions (e.g., Murphy and Zimmerman 1993; Pourciau 1993; Wells 2002; Cheng 2004).

In this study, I hypothesize that the impact of managerial retirement on discretionary accounting choices is contingent on managerial compensation. Specifically, the manager may have a Supplemental Executive Retirement Plan (SERP) under which the manager's pension is determined as a function of firm performance in the manager's final pre-retirement years. In this case, the manager has strong incentives to make accounting choices that increase firm short-term income and, therefore, amplify the value of his/her pension. Using a sample of CEOs of *Fortune* 1000 firms who retired between 1997 and 2006, I empirically investigate whether income-increasing accruals are associated with the horizon problem when CEO SERP benefits are contingent on firm performance. To verify the predicted relationships, I regress discretionary accruals on the interaction of the proxy for the horizon problem with the performance-contingency of CEO SERP. Consistent with most prior empirical studies, I find no evidence of income-increasing earnings management prior to CEO retirement in general. The results suggest that discretionary accruals are higher in CEO final pre-retirement years *only* when CEO SERP benefits are performance-contingent, in line with the hypothesis.

The study also investigates the market reaction to CEO retirement. Previous research suggests that voluntary CEO departures are generally perceived as orderly and well-planned events that do not lead to abnormal market returns (e.g., Huson et al. 2004). Consistent with this view, I find no evidence of long-term abnormal returns around CEO retirement announcements *on average*. However, a detailed analysis suggests that the market does not react to CEO retirement when the CEO has no SERP or has a SERP that is not contingent on firm performance, but does react to CEO retirement when CEO SERP benefits are performance-contingent. In the latter case, the post-retirement period is characterized by negative long-term abnormal returns. Furthermore, the significance and the magnitude of these returns are positively related to the size of discretionary accruals in the pre-retirement period. Taken together, the results suggest that the market does not unravel earnings management in the final pre-retirement years until after the retirement of the CEO, possibly because of an *ex ante* information asymmetry between shareholders and the CEO regarding the timing of CEO retirement. The results also imply that the equity of firms with performance-contingent CEO SERPs is temporarily overpriced in the period preceding CEO retirement.

Section II reviews previous studies on the impact of compensation contracts and managerial horizon on accounting discretion, describes the nature of SERPs, and develops the research hypothesis. Section III outlines the research methodology. Section IV discusses the results. Section V concludes.

II. HYPOTHESIS DEVELOPMENT

Literature Overview

There are numerous ways in which a manager can use judgment to make a particular accounting choice. Common examples include the choice of depreciation and inventory valuation methods and estimations of future events. Managers can also increase or decrease R&D, marketing, or any other business expense—operating decisions that ultimately impact firm financial reports. Over last two decades, earnings management has grown into a major research subject. This study builds upon and extends two streams of scientific literature: the first stream investigates the impact of managerial compensation on discretionary choices; the second investigates the impact of managerial horizon.

Numerous theoretical and empirical studies analyze relationships between managerial compensation contracts and discretionary choices. Watts and Zimmerman (1986) develop the bonus plan hypothesis, arguing that managers manipulate earnings in order to increase the amount of their bonus compensation. According to Healy (1985), firms that cap bonus awards are more likely to report income-deferring accruals when that cap is reached than similar firms with no bonus cap. However, later studies by Gaver et al. (1995) and Holthausen et al. (1995) reject the hypothesis that managers manipulate earnings in response to their bonus plans. On the other hand, Guidry et al. (1999) find that divisional managers are likely to defer income when the earnings target in their bonus plan will not be met and when they are entitled to maximum bonuses permitted under the plan. Murphy (2000) examines how the choice of external or internal performance standards affects firm performance and realized compensation, finding that firms with internal performance standards are characterized by lower variation in bonus payouts and are more likely to smooth earnings than firms with external performance standards.

Other compensation arrangements also appear to be linked to discretionary managerial choices. Bartov and Mohanram (2004) hypothesize that compensation contracts drive earnings management decisions and examine large stock option exercises by corporate insiders. They find that in the pre-exercise period, discretionary accruals are abnormally high, while in the post-exercise period, discretionary accruals are abnormally low. Both results are consistent with prior evidence that executives opportunistically time the option-grant date (Yermack 1997) and disclosures around option-grant dates (Aboody and Kasznik 2000).

A number of studies examine the relationship between managerial horizon and discretion. In theory, the manager who plans to leave the firm may lack incentives to act in the best interest of that firm. According to Smith and Watts (1982), the manager with a short horizon prefers projects with lower net present values but higher current accounting earnings to projects with higher net present values but lower current earnings. The horizon problem can also lead to earnings management, accounting fraud, and other consequences that are detrimental to various firm stakeholders. These activities should be especially pronounced when the manager intends to retire rather than join another firm (Gibbons and Murphy 1992), because a retiring manager faces fewer reputational concerns. However, despite theoretical predictions, empirical evidence on the impact of the horizon problem on discretionary accounting choices is scarce and inconclusive. According to Pourciau (1993), contrary to expectations, departing executives record accruals and write-offs that decrease earnings during their final year of tenure. DeAngelo (1988) finds that during proxy contests, incumbent managers exercise accounting discretion to improve reported earnings. On the other hand, using a sample of Australian firms, Wells (2002) finds little empirical evidence of income-increasing earnings management prior to CEO departures.

Furthermore, empirical findings on the association between managerial horizon and firm operating decisions are also mixed. Butler and Newman (1989) and Dechow and Sloan (1991) hypothesize that CEOs have incentives to reduce such discretionary expenditures as R&D and advertising in their final years to increase short-term earnings and corresponding bonuses. Dechow and Sloan (1991) find evidence supporting this argument. Gibbons and Murphy (1992) find that the growth rate of R&D intensity slows as CEOs approach retirement. According to Barker and Mueller (2002), CEO age is negatively associated with R&D expenditures. However, after controlling for endogenous CEO turnover and firm performance, Murphy and Zimmerman (1993) find little support for the impact of the horizon problem on R&D expenditures. Specifically, they find no evidence of managerial discretion in strongly performing firms when CEOs retire as part of the normal relay process—i.e., when executive change is an orderly and well-planned event that often concludes with the departing executive remaining a member of the board of directors (Vancil 1987). In addition, Cheng (2004) finds no association between CEO turnover and R&D expenditures, suggesting that Dechow and Sloan's (1991) different result could be attributable to a different time period.

In sum, in general, existing empirical literature suggests that managers make accounting choices to increase their compensation but does not provide conclusive evidence that the likelihood and/or the magnitude of earnings management are greater in years immediately prior to managerial retirement. However, as the following discussion suggests, some managers have stronger incentives to manage earnings in their final pre-retirement years than others due to the nature of their retirement plans.

Supplemental Executive Retirement Plans

SERPs exist in most developed countries due to regulatory limitations on the final retirement income under regular pension plans. In the United States, the Internal Revenue Code caps the annual compensation used for determining pension benefits under a qualified plan (i.e., pensionable earnings); in 2009, the cap is set at $235,000.[1] Consider the CEO who retires with a $2,000,000 pre-retirement base salary, 35 years of pensionable service, and a pension plan that calls for 2 percent of the final base salary multiplied by the number of years of pensionable service to be paid to the CEO annually upon retirement. In this case, the CEO's pension is limited to $164,500.[2] Since seven-digit executive compensation figures are common on the market, especially in larger firms, a regular pension represents a modest fraction of pre-retirement income for many executives.[3]

A SERP is a non-contributory pension plan funded by the employer, which serves to increase an executive's post-retirement income beyond the regular pension limit. Under a SERP, a firm makes ongoing payments to a retired executive on top of the regular pension until his/her death or even until the death of the surviving spouse. Research shows that SERPs are common and sizable. Sundaram and Yermack (2007) investigate a sample of *Fortune* 500 firms and find that 78 percent of CEOs in the sample have supplemental retirement plans. The mean value of CEO annual pension increment (i.e., the amount by which the actuarial value of a SERP increases over a given year during CEO employment) is $0.9 million, which makes it comparable to CEO base salary or cash bonus. Kalyta and Magnan (2008) obtain similar results for a limited sample of largest Canadian firms. The latter study also finds that the pension increment is substantially higher in the final year prior to CEO retirement than in any other year, suggesting a possibility of income-increasing earnings management. Finally, Kalyta (2009) argues that managerial opportunism is inversely related to the transparency of managerial pay and provides some evidence that CEO SERP benefits—historically the least-transparent compensation component—are associated with the extraction of rents and reductions in R&D expenditures in largest Canadian firms. While these findings do not imply that SERPs lead to earnings management, they further legitimize an inquiry into the association between CEO SERP benefits and discretionary accounting choices.

Annual SERP benefits are typically calculated as a certain percentage (multiplier) of pensionable earnings *times* number of years of pensionable service *minus* capped regular pension benefits. However, specific SERP components and provisions vary significantly from one CEO to another. Specifically, all CEOs can be classified into three groups: (1) CEOs without SERP arrangements; (2) CEOs with SERP arrangements in which pensionable earnings consist of base salary only; and (3) CEOs with SERP arrangements in which pensionable earnings are determined as a function of salary, performance bonus, and, sometimes, other performance-based compensation components. An increase in accounting earnings is most beneficial to CEOs from the latter group, since an increase in performance-based compensation amplifies the value of the CEO pension plan. Under existing SERPs, pensionable earnings are usually determined in final several years immediately prior to retirement (hereafter, determination years), so the beneficial impact of accounting earnings on CEO compensation is especially pronounced in these years. The following example illustrates the potential magnitude of the impact of performance-based compensation on CEO SERP benefits. Consider a retiring CEO, with 20 years of credited service, aged 60, whose annual pension is determined by the product of cash compensation (base salary *plus* annual performance bonus) in the final year prior to retirement, years of credited service, and a multiplier of 2 percent. If the salary of the CEO in the final year prior to retirement is \$1,000,000 and the bonus is \$0, then the CEO's annual pension will be \$400,000 and the actuarial value of the pension plan will be \$5,637,378, assuming a discount rate of 0.05, current death probability tables, and no surviving spouse. However, if the salary of the CEO in the final year prior to retirement is \$1,000,000 and the CEO also receives a \$1,000,000 bonus, then the CEO's annual pension will be \$800,000 and the actuarial value of the pension plan will be \$11,275,156. In other words, \$1,000,000 of bonus compensation will actually increase the wealth of the CEO by \$6,637,378. Neither CEOs without SERPs nor CEOs with SERP arrangements in which pensionable earnings consist of base salary only can enjoy such significant increases in their wealth and retirement security. Consequently, CEOs with performance-contingent SERPs have strong incentives to make accounting choices that increase firm short-term income in determination years and, therefore, amplify the value of their pensions. This leads to the following prediction:

---

[1] Sections 401(a)(17), 404(l), 408(k)(3)(C), and 408(k)(6)(D)(ii) of the Internal Revenue Code. Each year, the cap is modified to reflect changes in the cost of living.

[2] 35 × 0.02 × $235,000.

[3] The mean annual CEO cash compensation in a sample of *Fortune* 500 firms during the 1996–2002 period is $2.3 million (Sundaram and Yermack 2007).

> **Hypothesis:** Discretionary accruals are positively associated with the horizon problem when CEO SERP benefits are contingent on firm performance.

By analyzing associations between CEO SERP benefits and discretionary accruals, the study makes an important contribution to several streams of academic research. First, I investigate a previously unexplored dimension of the relationship between the horizon problem and discretionary accounting choices. Empirical evidence in this research area is mixed and inconclusive. Second, investigating the links between CEO SERP benefits and accounting discretion contributes to the stream of academic research on the role of executive compensation. Existing studies in this research area provide evidence that specific executive compensation components do impact accounting choices. At the same time, implications of CEO SERP benefits remain largely unexplored, even though SERP benefits represent a unique post-retirement compensation arrangement with incentives that are distinct from other executive pay components.

III. METHODOLOGY

McNichols and Wilson (1988) outline a general framework for classifying accounting accruals into normal (non-discretionary) and abnormal (discretionary) components:

$$TACC = NDACC + DACC \tag{1}$$

where:

$TACC$ = total accruals;
$NDACC$ = non-discretionary accruals; and
$DACC$ = discretionary accruals.

Discretionary accruals are considered to be the outcome of managerial opportunistic choices. To calculate discretionary accruals, I first measure total accruals using the cash-flow method suggested by Hribar and Collins (2002):

$$TACC_{it} = EBXI_{it} - CFCO_{it} \tag{2}$$

where $TACC$ is as defined earlier and:

i = firm subscript;
t = year subscript;
$EBXI$ = earnings before extraordinary items and discontinued operations; and
$CFCO$ = cash flows from continuing operations.

To measure discretionary accruals, I rely on an extended time-series version of the widely used modified Jones model (Dechow et al. 1995):

$$\begin{aligned} TACC_{it}/A_{it-1} = {} & \beta_0/A_{it-1} + \beta_1(\Delta REV_{it} - \Delta REC_{it})/A_{it-1} + \beta_2 PPE_{it}/A_{it-1} \\ & + \beta_3 CFCO_{it}/A_{it-1} + \varepsilon \end{aligned} \tag{3}$$

where:

A = total assets;
ΔREV = change in sales revenue;
ΔREC = change in accounts receivable; and
PPE = property, plant, and equipment.

Discretionary accruals are then defined as the residual from Equation (3):

$$\begin{aligned} DACC_{it} = {} & TACC_{it}/A_{it-1} - [\beta_0/A_{it-1} + \beta_1(\Delta REV_{it} - \Delta REV_{it})/A_{it-1} \\ & + \beta_2 PPE_{it}/A_{it-1} + \beta_3 CFCO_{it}/A_{it-1}]. \end{aligned} \tag{4}$$

Finally, to test the hypothesis, I estimate the following empirical model:

$$\begin{aligned} DACC_{it} = {} & \beta_0 + \beta_1 HORIZON_{it} * PERFSERP_{it} + \beta_2 HORIZON_{it} \\ & + \beta_3 PERFSERP_{it} + \beta_4 LSIZE_{it} + \beta_5 LEV_{it} + \beta_6 GROWTH_{it} \\ & + \beta_7 PERF_{it-1} + \beta_8 LCFCO_{it} + \beta_9 DEVREV_{it} \\ & + \beta_{10} DEVCFCO_{it} + \Sigma\beta_j INDUSTRY_j + \varepsilon \end{aligned} \tag{5}$$

where:

HORIZON = proxy for the horizon problem, which in turn comprises the following indicator variables: $HORIZON^{DET}$ is equal to 1 if *t* is a determination year, and 0 otherwise; and $HORIZON^{FIN}$ is equal to 1 if *t* is the final year prior to CEO retirement, and 0 otherwise;
PERFSERP = indicator variable equal to 1 if CEO pensionable earnings include performance-based compensation, and 0 if CEO pensionable earnings include base salary only or if CEO does not have a SERP;
LSIZE = natural logarithm of total assets;
LEV = debt-to-assets ratio;
GROWTH = book-to-market ratio;
PERF = return on assets;
LCFCO = natural logarithm of cash flows from continuing operations;
DEVREV = standard deviation of revenue deflated by lagged total assets over the current and prior four years;
DEVCFCO = standard deviation of cash flows from continuing operations deflated by lagged total assets over the current and prior four years; and
INDUSTRY = indicator variables for industry classification.

The *HORIZON* variables partition the sample into observations affected by the horizon problem and all other observations. The *PERFSERP* variable partitions the sample into observations with performance-contingent SERPs and all other observations. The variable of interest is the interaction between the horizon problem and the performance-contingency of CEO SERP benefits, *HORIZON* * *PERFSERP*. The interaction can be interpreted as a dichotomous variable that takes the value of 1 in observations with hypothesized earnings management and the value of 0 in observations with no hypothesized earnings management. According to the hypothesis, the β_1 coefficient should be positive and significant; i.e., earnings management is expected only when CEOs with short horizons have performance-contingent SERPs. To clarify whether the horizon problem arises in the final year prior to CEO retirement or when the determination period begins, I employ two alternative proxies for the horizon problem: $HORIZON^{DET}$ and $HORIZON^{FIN}$. In the first case, the horizon problem spans over a longer period of time: the impact on discretionary accruals is hypothesized in all determination years. In the second case, the horizon problem affects a shorter period: income-increasing earnings management is hypothesized only in the final year prior to CEO retirement. I run separate regressions for each proxy.

Since multiple annual observations for a particular CEO are expected to be correlated, the regressions rely on standard panel data models with Huber-White standard errors that account for CEO-specific effects. Consistent with prior empirical studies, the model controls for common firm-level factors that may affect discretionary accounting choices: performance (e.g., Klein 2002; Chung and Kallapur 2003; Butler et al. 2004), growth opportunities (Klein 2002; Menon and Williams 2004; Butler et al. 2004), leverage (DeFond and Jiambalvo 1994; Becker et al. 1998; Kim et al. 2003; Butler et al. 2004), and cash flows from operations (Becker et al. 1998; Chung and Kallapur 2003; Frankel et al. 2002; Kim et al. 2003; Menon and Williams 2004). I use ROA as a proxy for firm performance, debt-assets ratio as a proxy for leverage, and book-to-market ratio as a proxy for growth opportunities. Although the measure of discretionary accruals is deflated by total assets, an association between *DACC* and firm size may still exist (e.g., Kim et al. 2003; Menon and Williams 2004). Therefore, I include the natural logarithm of firm assets in earnings management tests as another control variable. Following Hribar and Nichols (2007), the model controls for the volatility of firm revenue and the volatility of operating cash flows over a five-year period.[4] Finally, I include industry-specific fixed effects in the model. Given the size of the sample, I employ an updated version of the industry classification from Barth et al. (1999) instead of more detailed two- or three-digit SIC classifications.

IV. RESULTS

Data

The study uses the sample of firms that made the *Fortune* 1000 list in 2006, with non-public firms and firms in the financial services industry excluded from the analysis, and hence is comprised of relatively large U.S. firms. Empirical evidence shows a positive association between CEO cash compensation and firm size (e.g., Jensen and Murphy 1990; Lambert et al. 1991; Craighead et al. 2004; Kalyta and Magnan 2008). Higher pre-retirement cash compensation leads to a greater disproportion between pensionable earnings and regular pension and, therefore, to a greater role of the SERP in the compensation contract.[5] Consequently, the impact of CEO SERP on discretionary accounting choices is expected to be especially pronounced in larger firms. I collect data for the period between 1997 and 2006. Information on CEO departures and retirement plans is collected directly from annual proxy statements (i.e., Schedule 14A). Financial data are retrieved from Compustat and corporate financial reports. To account for inflation, I convert all monetary values into 2006 dollars using historical Consumer Price Index data.

Sample Selection and Descriptive Statistics

The initial sample contains 2,013 CEOs. Current CEOs of *Fortune* 1000 firms are excluded from the analysis. This reduces the sample to 1,137 CEOs who left their firms between 1997 and 2006. Gibbons and Murphy (1992), among others, suggest that the horizon problem should be especially pronounced when the CEO approaches retirement as opposed to leaving the firm and staying on the job market. In the former case, the CEO's concerns about the discipline from managerial labor markets become less significant or disappear. Murphy and Zimmerman (1993) argue that, in theory, retirements are more likely to result in income-increasing earnings management in pre-departure years than unplanned departures, since the CEO can anticipate retirement and plan discretionary choices in advance. To eliminate observations corresponding to CEO departures other than retirements, I classify departure as a retirement if the CEO's age at departure is 60 or higher.[6] When CEO age is not disclosed in annual proxy statements, I retrieve it via the *Wall Street Journal* Index, Lexis-Nexis databases, and Google. Although it is probable that some CEOs stay on the job market after turning 60 while some CEOs retire before reaching that age, imposing this criterion reduces the number of observations with hypothesized earnings management in which the horizon problem is less likely to be significant.

---

[4] Alternatively, I use ROE as a proxy for firm accounting performance, debt-equity ratio as a proxy for firm leverage, the natural logarithm of firm revenue as a proxy for firm size, and three-year volatilities of firm revenue and cash-flows from continuing operations instead of five-year volatilities. Results are not qualitatively affected.

[5] Consider two retiring CEOs with SERPs, with 35 years of credited service, whose annual pension is to be determined by the product of pre-retirement cash compensation, years of credited service, and a multiplier of 2 percent. Pre-retirement cash compensation of CEO A is $250,000. Pre-retirement cash compensation of CEO B is $2,500,000. In that case, CEO A is entitled to annual pension of $175,000, of which 94 percent (35 × 0.02 × $235,000 = $164,500) is from the regular pension, and the remaining 6 percent is from the SERP. CEO B is entitled to annual pension of $1,750,000, of which only 9.4 percent ($164,500) is from the regular pension, and 90.6 percent is from the SERP.

[6] Alternatively, I perform the analysis assuming the age of retirement is 63. The choice of the cut-off point does not qualitatively affect the main results.

Also, if the horizon problem affects accounting choices in the final years prior to CEO retirement, the next CEO will have limited means to make income-increasing discretionary choices due to accrual reversals. To alleviate the potential impact of the horizon problem associated with the departure of the previous CEO, I impose another criterion: a retired CEO must have held the position for more than two full years. This reduces the sample to 503 retired CEOs. Finally, eliminating observations with missing firm-level data required for the analysis reduces the sample to 388 retired CEOs. Table 1 summarizes the sample selection process (Panel A), and provides information on the industry composition of the sample (Panel B) and its structure (Panel C). Most CEOs have supplemental retirement plans (70.4 percent). In this group, 172 CEOs have performance-contingent SERPs in which the pension is determined as a function of performance-based pay.[7] In total, there are 2,011 observations, or CEO-years, in the sample, including 880 determination years and 273 final years. The average determination period is 3.22 years. Depending on the proxy for the horizon problem ($HORIZON^{DET}$ or $HORIZON^{FIN}$), the sample contains 518 (25.8 percent) or 170 (8.5 percent) observations with hypothesized income-increasing earnings management.

Table 2 provides descriptive statistics for firm variables in the final sample (Panel A) and reports the correlation matrix (Panel B). Most firms in the sample are profitable. The median firm earns a 12.8 percent return on its assets, has a 42.5 percent leverage ratio, and a 37.8 percent book-to-market ratio. In general, these results are consistent across observations with performance-contingent CEO SERPs ($PERFSERP = 1$) and observations with either no SERPs or SERPs that are not contingent on firm performance ($PERFSERP = 0$). Most correlation coefficients are significant; however, only the correlation between firm size and cash flows from continuing operations exceeds 0.4. Overall, correlations appear to be intuitive and consistent with the prior literature.

Earnings Management Tests

To perform a univariate analysis of the hypothesis, I partition the sample into observations with and without hypothesized income-increasing discretionary choices and compare measures of discretionary accruals across the two subsamples. Table 3, Panel A, reports the results. The analysis considers two alternative definitions of the horizon problem. In the first case, I assume a broader definition of the horizon problem: income-increasing discretionary choices are hypothesized when CEOs with performance-contingent SERPs are in their determinations years (i.e., observations that satisfy $HORIZON^{DET} * PERFSERP = 1$). The t-tests indicate that observations with income-increasing incentives are characterized by significantly higher discretionary accruals. In the second case, I assume a narrower definition of the horizon problem. Income-increasing discretionary choices are hypothesized when CEOs with performance-contingent SERPs are in their final year prior to retirement (i.e., observations that satisfy $HORIZON^{FIN} * PERFSERP = 1$). The t-tests find a marginally significant difference in discretionary accruals between the two subsamples. This can be explained by the fact that observations corresponding to determination years *other than the final year* fall into the subsample in which income-increasing discretionary choices are not hypothesized. As an additional test (untabulated), I exclude observations

---

[7] In some SERPs, pensionable earnings include base salary and *target* cash bonus. There is no direct impact of income-increasing accounting choices on the value of such SERPs in a given year. Consequently, SERPs in which pensionable earnings include base salary and target cash bonus are classified as SERPs not contingent on firm performance.

TABLE 1
Sample Selection, Industry Composition, and SERP Statistics

Panel A: Sample Selection

Observation	n
Total CEOs	2,013
Less: Actual CEOs	(876)
Departed CEOs	1,137
Less: Departures other than retirements and short tenures (two years or less)	(634)
Retired CEOs	503
Less: Missing data	(115)
Retired CEOs, final sample	388

Panel B: Industry Composition

#	Industry	SIC Codes	n	%
01	Agriculture-Food	000–900, 2000–2111	24	6.2
02	Mining-Petroleum	1000–1400, 2911–2990	29	7.5
03	Construction-Real Estate	1520–1731, 6500–6799	16	4.1
04	Textiles-Paper-Printing	2200–2790	18	4.6
05	Chemicals-Pharmaceuticals	2800–2891	25	6.4
06	Manufacturing	3000–3569, 3580–3590, 3690–3990	49	12.6
07	Computers-Electronics-Communications	3570–3579, 3600–3679, 4812–4899, 7370–7377	46	11.9
08	Transportation-Services	4011–4731, 7000–7363, 7380–8744	64	16.5
09	Utilities	4900–4991	32	8.2
10	Retail	5000–5990	50	12.9
11	Finance	6021–6282	0	0.0
12	Insurance	6311–6411	35	9.0
	Total		388	100.0

Panel C: SERP Statistics

Observation	n	%
Retirement Arrangements		
CEOs with SERPs	273	70.4
CEOs without SERPs	115	29.6
Total	388	100.0
SERP Structure		
SERPs contingent on firm performance	172	63.0
SERPs *not* contingent on firm performance	101	37.0
Total	273	100.0
Determination Period		
Determination years	880	43.8
Other years	1,131	56.2
Total	2,011	100.0
Average determination years per CEO	3.22	NA

(*continued on next page*)

TABLE 1 (continued)

Observation	n	%
Final Years		
Final years	273	13.6
Other years	1,738	86.4
Total	2,011	100.0
Observations with Hypothesized Earnings Management ($HORIZON^{DET}$)		
SERPs contingent on performance * Determination years	518	25.8
Other observations	1,493	74.2
Total	2,011	100.0
Observations with Hypothesized Earnings Management ($HORIZON^{FIN}$)		
SERPs contingent on performance * Final years	170	8.5
Other observations	1,841	91.5
Total	2,011	100.0

This table provides information on the sample selection procedure (Panel A), the industry composition of the sample (Panel B), and the structure of retirement arrangements of CEOs of *Fortune* 1000 firms who retired between 1997 and 2006 (Panel C). Percentages may not add up to 100 percent due to rounding. Firms in the financial services industry are excluded from analysis.

corresponding to determination years other than the final year and compare discretionary accruals across the two subsamples. The difference between abnormal accruals becomes significant at the 5 percent level. In sum, the results suggest that the horizon problem arises earlier than the final year prior to CEO retirement. More importantly, the results are consistent with the prediction that discretionary accruals are higher when CEOs with performance-contingent SERPs are in their determination years.

Figure 1 depicts evidence of income-increasing discretionary accruals in firms with performance-contingent CEO SERPs in the period preceding CEO retirement. The graph, centered on the year of CEO retirement, suggests substantial increases in discretionary accruals in Years −4 and −3. These increases are consistent with the structure of CEO SERPs: an average period during which CEO pension is determined lasts 3.22 years (Table 1, Panel C). The value of discretionary accruals peaks in Year −2 and declines, but remains positive in the last year prior to CEO retirement.[8] Also, consistent with the expectation of accrual reversals, the period following CEO retirement is characterized by negative discretionary accruals. The measure of discretionary accruals is at its minimum in Year 2 following CEO retirement.

To further investigate the hypothesis, I estimate Equation (5) using two separate OLS regressions, with *DACC* as the dependent variable and either $HORIZON^{DET}$ or $HORIZON^{FIN}$ as the proxy for the horizon problem. The explanatory variable of interest is the interaction between the horizon problem and the performance-contingency of CEO supplemental pension arrangements, *HORIZON* * *PERFSERP*. To mitigate any influence from outliers, I

---

[8] There are several potential explanations of this decline. First, due to eventual accrual reversals, the CEO may have limited means to increase firm short-term earnings after several years of abnormally high discretionary accruals. Second, the relationship between performance-based CEO compensation and lagged firm performance may be stronger than the relationship between performance-based CEO compensation and contemporaneous firm performance. In other words, firm performance in the last year prior to CEO retirement may have little effect on the magnitude of CEO compensation. Third, firm stakeholders may anticipate CEO retirement in the last year prior to CEO retirement and employ governance and other control mechanisms that alleviate the horizon problem.

TABLE 2
Descriptive Statistics

Panel A: Firm Variables[a]

	Median			Other Statistics (full sample)		
		PERFSERP =				
Variable	Full Sample	1	0	Mean	Standard Deviation	n
LSIZE	21.94	21.98	21.60	22.86	23.89	2,011
LEV	0.425	0.429	0.412	0.480	0.212	2,011
GROWTH	0.378	0.360	0.386	0.476	0.261	2,011
PERF	0.128	0.130	0.128	0.134	0.083	2,011
LCFCO	19.66	19.76	19.53	20.15	20.50	2,011
DEVREV	0.270	0.285	0.261	0.362	0.343	2,011
DEVCFCO	0.081	0.080	0.084	0.136	0.140	2,011

Panel B: Correlations among Firm Variables[b]

Variable	*LSIZE*	*LEV*	*GROWTH*	*PERF*	*LCFCO*	*DEVREV*	*DEVCFCO*
LSIZE		0.22 (0.00)	−0.01 (0.74)	0.02 (0.88)	0.52 (0.00)	−0.24 (0.00)	−0.36 (0.00)
LEV	0.27 (0.00)		0.31 (0.00)	−0.25 (0.00)	0.27 (0.00)	−0.06 (0.22)	−0.09 (0.10)
GROWTH	−0.05 (0.28)	0.34 (0.00)		−0.19 (0.00)	−0.07 (0.27)	−0.02 (0.75)	−0.04 (0.51)
PERF	0.03 (0.85)	−0.17 (0.00)	−0.21 (0.00)		0.35 (0.00)	0.01 (0.86)	0.04 (0.49)
LCFCO	0.45 (0.00)	0.29 (0.00)	−0.10 (0.14)	0.35 (0.00)		−0.13 (0.09)	−0.17 (0.01)
DEVREV	−0.20 (0.00)	−0.02 (0.73)	−0.01 (0.91)	0.03 (0.60)	0.14 (0.05)		0.35 (0.00)
DEVCFCO	−0.24 (0.00)	−0.07 (0.24)	−0.05 (0.42)	0.06 (0.38)	0.26 (0.00)	0.28 (0.00)	

---

[a] Panel A reports descriptive statistics and correlation coefficients for the sample of CEOs of *Fortune* 1000 firms who retired between 1997 and 2006.

[b] Pearson (Spearman) correlations are above (below) the diagonal. p-values are reported in parentheses.

Variable Definitions:
LSIZE = natural logarithm of total assets;
LEV = debt-to-assets ratio;
GROWTH = book-to-market ratio;
PERF = return on assets;
LCFCO = natural logarithm of cash flows from continuing operations;
DEVREV = standard deviation of revenue deflated by lagged total assets over the current and prior four years; and
DEVCFCO = standard deviation of cash flows from continuing operations deflated by lagged total assets over the current and prior four years.

TABLE 3
Analysis of Discretionary Accruals[a]

Panel A: Univariate Analysis[b]

Variable	*HORIZON * PERFSERP =* **1**	*HORIZON * PERFSERP =* **0**	**Difference**
$HORIZON^{DET}$			
DACC	0.0020	−0.0008	0.0028 (0.003)
n	518	1,493	
$HORIZON^{FIN}$			
DACC	0.0008	−0.0000	0.0008 (0.086)
n	170	1,841	

Panel B: Multivariate Analysis[c]

	Unadjusted Sample		**Matched Sample**	
Variable	$HORIZON^{DET}$	$HORIZON^{FIN}$	$HORIZON^{DET}$	$HORIZON^{FIN}$
*HORIZON * PERFSERP*	0.0211 (0.008)	0.0108 (0.090)	0.0286 (0.018)	0.0278 (0.077)
HORIZON	0.0024 (0.120)	0.0010 (0.119)	0.0040 (0.125)	0.0065 (0.100)
PERFSERP	0.0153 (0.194)	0.0087 (0.216)	0.0193 (0.184)	0.0060 (0.332)
LSIZE	−0.0070 (0.000)	−0.0095 (0.000)	−0.0090 (0.000)	−0.0097 (0.001)
LEV	0.0199 (0.077)	0.0313 (0.073)	−0.0054 (0.440)	0.0103 (0.292)
GROWTH	0.0067 (0.008)	0.0055 (0.008)	0.0019 (0.051)	0.0022 (0.085)
PERF	0.4925 (0.001)	0.4844 (0.002)	0.5337 (0.002)	0.5715 (0.001)
LCFCO	−0.2480 (0.005)	−0.2362 (0.005)	−0.2981 (0.004)	−0.3374 (0.002)
DEVREV	0.0070 (0.122)	0.0092 (0.101)	0.0062 (0.226)	−0.0021 (0.514)
DEVCFCO	0.0658 (0.076)	0.0561 (0.085)	−0.0152 (0.367)	−0.0119 (0.530)
n	2,011	2,011	1,036	340
Adjusted R^2	0.286	0.283	0.231	0.196

---

[a] This table displays univariate and multivariate analysis of firm discretionary accruals for the sample of CEOs of *Fortune* 1000 firms who retired between 1997 and 2006. Multivariate models are estimated using OLS with robust Huber-White standard errors and control for industry effects (not tabulated). p-values for two-tailed t-tests are reported in parentheses

[b] Panel A reports descriptive statistics for observations with and without hypothesized income-increasing discretionary choices.

[c] Panel B displays parameter estimates of the determinants of unsigned discretionary accruals, estimated from Equation (4).

Variable Definitions:

$HORIZON^{DET}$ = an indicator variable equal to 1 if the observation corresponds to a determination year;

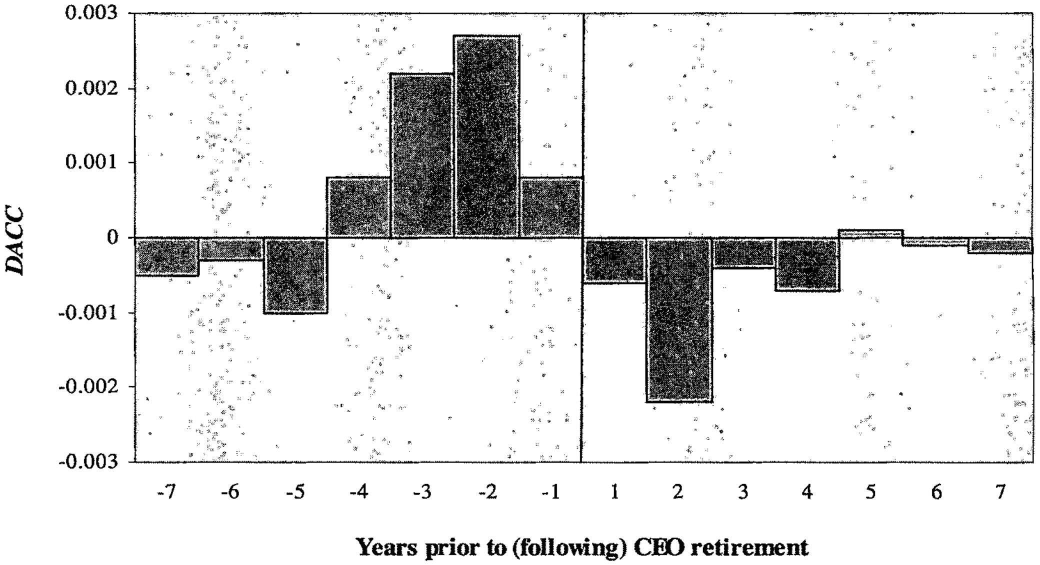
$HORIZON^{FIN}$ = an indicator variable equal to 1 if the observation corresponds to the final year prior to CEO retirement; and

PERFSERP = dichotomous variable equal to 1 if CEO pensionable earnings include performance-based pay.

Other variables are described in Table 2. All variables are winsorized at the 1 percent level.

winsorize all variables at the 1 percent level. The analysis supports the univariate results: firms in which CEOs with performance-contingent SERPs are in their determinations years are characterized by larger abnormal accruals. Results are reported in the second column of Table 3, Panel B. The coefficient on the $HORIZON^{DET}$ * *PERFSERP* interaction is positive and significant (0.0211; $p < 0.01$). When considered separately, neither the horizon problem nor the performance-contingency of CEO SERP have a significant association with

FIGURE 1
Discretionary Accruals in Firms with Hypothesized Earnings Management



This figure illustrates time-series data on discretionary accruals in firms with performance-contingent CEO SERPs. The horizontal axis is centered on the year of CEO retirement.

abnormal accruals. In other words, a positive association between the horizon problem and income-increasing accounting choices exists only when SERP benefits of the CEO are contingent on firm performance.[9]

The association between accounting discretion and *HORIZON* * *PERFSERP* becomes marginally significant when income-increasing discretionary choices are restricted to the final year prior to CEO retirement. Results are reported in the third column of Table 3, Panel B. The coefficient on the $HORIZON^{FIN}$ * *PERFSERP* interaction is 0.0108 ($p < 0.10$).

---

[9] An implicit assumption in the study is that the effect of firm performance on performance-based pay is greater than the effect of firm performance on base salary. While the assumption is reasonable due to the very nature of compensation contracts, it is possible that in some cases the pay deemed to be performance-based (e.g., cash bonus) is a function of a measure over which the CEO has no or little control. On the other hand, it is possible that in some cases, the salary increases beyond a "normal" annual increase as a result of firm performance while "performance-based" compensation components remain unaffected. However, the two events are unlikely to be common and would bias *against* the results presented in this manuscript. Nevertheless, to corroborate the findings, I perform a simple albeit imperfect pay-for-performance analysis. First, for each firm with ten years of data available, I run the regression of annual changes in CEO base salary on lagged annual changes in firm ROA, and the regression of annual changes in CEO actual cash bonus on lagged annual changes in firm ROA. I remove from the analysis the firms for which the R^2 in the base salary regression is higher than the R^2 in the cash bonus regression. I then rank the sample of retired CEOs with performance-contingent SERPs according to the R^2 in the cash bonus regression. I remove observations in the lower 50th percentile. In other words, the remaining observations with hypothesized income-increasing discretionary choices are limited to observations with highest (50 percent) bonus-to-performance contingency and exclude cases in which the cash bonus is less contingent on firm performance than the base salary is. In the final step, I combine the remaining observations with the control sample (i.e., observations for which *HORIZON* * *PERFSERP* = 0) and re-estimate Equation (5). The main result remains unaffected; coefficients on the *HORIZON* * *PERFSERP* interaction are positive and significant in explaining discretionary accruals, thereby corroborating the hypothesis.

As an additional test (untabulated), I estimate the model without observations in determination years other than the final year. When the sample is limited to observations corresponding to the final year prior to CEO retirement and observations in non-determination years, the coefficient on the $HORIZON^{FIN} * PERFSERP$ interaction is positive and significant at the 1 percent level. As in the main model, when considered separately, neither the horizon problem nor the performance-contingency of CEO SERP are significantly associated with abnormal accruals.

Kothari et al. (2005) suggest that using performance-based matching techniques enhances the reliability of earnings management tests. To triangulate the results, I match observations with hypothesized income-increasing discretionary accruals (i.e., $HORIZON * PERFSERP = 1$) with observations with no hypothesized income-increasing discretionary accruals (i.e., $HORIZON * PERFSERP = 0$) based on the industry classification and the closest ROA in the current year. This approach reduces the sample size to 1,036 observations in the test with $HORIZON^{DET}$ and 340 observations in the test with $HORIZON^{FIN}$. Results are reported in the last two columns of Table 3, Panel B. Overall, the findings remain qualitatively unchanged: a positive association between income-increasing discretionary choices and the horizon problem exists only when CEO SERP is contingent on firm performance.

In sum, results provide empirical support for the hypothesized relationship. Consistent with prior studies, I find no association between accounting discretion and the horizon problem in general. However, a positive association between income-increasing discretionary choices and the horizon problem does exist when the value of CEO supplemental pension is contingent on firm performance in CEO pre-retirement years. Results also indicate that the horizon problem arises earlier than in the final year prior to CEO retirement—when the period in which CEO pensionable earnings are determined begins.[10]

Market Reaction to CEO Retirement

Although the market may have expectations about management turnover, the CEO is in a better position to predict the timing of his/her retirement than any other individual. In other words, there exists an information asymmetry between shareholders and the CEO regarding the timing of CEO retirement. Hence, any increases in discretionary accruals in the final years prior to CEO retirement may not be unraveled by the market until after the retirement announcement, causing a temporary overpricing of firm equity. As a result, I expect the market to react more negatively (or less positively) to CEO retirement if the CEO has a performance-contingent SERP. To investigate this prediction, I examine abnormal stock returns around CEO retirement announcements. Although many empirical studies investigate stock market reactions to CEO changes (e.g., Weisbach 1988; Denis and Denis 1995; Huson et al. 2004), no prior research accounts for the structure of CEO retirement arrangements.

---

[10] I perform several additional tests to analyze the sensitivity of the findings. First, estimating total accruals with the balance sheet approach instead of the cash-flow approach does not affect the results qualitatively. Under the balance sheet approach, total accruals are calculated as change in current assets *minus* change in current liabilities *minus* change in cash and cash equivalents *plus* change in short-term debt *minus* change in depreciation and amortization expense, with the result deflated by lagged total assets to control for scale differences. Second, the results remain qualitatively unchanged when the Jones (1991) model is used to estimate abnormal accruals, instead of the modified Jones model. Under the Jones model, abnormal accruals are estimated as the residual from the regression of current accruals on the reciprocal of total assets, change in revenue and property, plant, and equipment, with all variables deflated by lagged total assets. Third, the results remain qualitatively unchanged when the proxy for size is not included in the main regression. Fourth, the results are unaffected if abnormal current accruals are used as the dependent variable. For this purpose, current accruals are defined as the change in the non-cash working capital. The abnormal component is the residual from the regression of current accruals on the reciprocal of total assets, change in revenue *minus* change in receivables, and cash flows from continuing operations, with all variables deflated by lagged total assets. Fifth, the results are not affected by the inclusion of book-to-market ratio as an additional explanatory variable in the Modified Jones model. Finally, using $PERF_{it}$ instead of $PERF_{it-1}$ as an explanatory variable in Equation (5) does not impact the results. In all alternative models, the coefficient on the $HORIZON * PERFSERP$ interaction remains significant when $HORIZON = HORIZON^{DET}$. When $HORIZON = HORIZON^{FIN}$, the coefficient on the $HORIZON * PERFSERP$ interaction becomes significant at the 5 percent level when the Jones model is used and remains marginally significant in all other alternative models.

To determine the impact of the announcement of CEO retirement on the value of firm equity, I follow the event study methodology employed by Huson et al. (2004). Huson et al. (2004) assess stock returns in three-year periods around CEO turnover events, finding no evidence of abnormal returns either before or after voluntary CEO departures—a subsample of turnovers that includes, but does not focus specifically on, retirements. For each CEO retirement in the sample, I use the *Wall Street Journal* Index, Lexis-Nexis historical databases, Google, and corporate proxy statements to determine the calendar month in which the retirement is announced. I then compute abnormal stock returns using calendar-time portfolios. For each month t, all firms in which the CEO will retire within the next 36 months are included in that month's pre-retirement portfolio. All firms in which the CEO has retired within the previous 36 months are included in the post-retirement portfolio. Observations that can be classified as both pre- and post-retirement are not included in either portfolio. Following Huson et al. (2004), I also exclude from the analysis any month with fewer than ten event-firms in the portfolio. I estimate equally weighted and value-weighted portfolio returns for each month and regress abnormal event portfolio returns on the three Fama and French (1993) factors:[11]

$$(rp_t - rf_t) = \beta_0 + \beta_1(rm_t - rf_t) + \beta_2 HL_t + \beta_3 SB_t + \varepsilon \tag{6}$$

where:

rp = return on event portfolio (pre- or post-retirement);
rf = one-month Treasury bill rate;
rm = return on all NYSE, AMEX, and NASDAQ stocks in CRSP database;
HL = return difference between portfolios of high and low book-to-market stocks; and
SB = return difference between portfolios of small and big capitalization stocks.

Under the null hypothesis of no abnormal performance, the intercept β_0 measures the average monthly abnormal return on the portfolio of pre- or post-retirement firms.[12] Following Mitchell and Stafford (2000) and Huson et al. (2004), I adjust the intercept for expected abnormal returns attributable to the portfolio composition. I calculate the adjusted intercept as the difference between β_0 and the mean intercept from 1,000 three-factor models of returns on random non-event portfolios with the same characteristics (size, book-to-market ratio, and calendar-time frequency) as the event portfolios.[13]

---

[11] Data are obtained from Kenneth French's website (http://mba.tuck.dartmouth.edu/pages/faculty/ken.french).

[12] I focus on long-term rather than short-term abnormal stock returns around CEO retirement announcements for several reasons. First, the accuracy of short-term stock analysis is highly sensitive to the precise knowledge of the event date. In many cases, it is impossible to identify the exact date when an announcement of CEO retirement becomes public (e.g., the CEO announces his/her retirement during a board meeting on Friday, the firm posts the press-release on its website on Monday, but the *Wall Street Journal* covers the story on Wednesday). Second, analyzing long-term returns allows assessing whether the market reacts differently to various CEO retirement arrangements in the years preceding CEO retirement, especially if these years are characterized by high discretionary accruals. Finally, a long-term approach allows comparing the results to those reported in Huson et al. (2004)—one of the most recent studies that analyze stock returns around CEO turnovers in largest U.S. firms.

Table 4 reports abnormal stock returns around CEO retirement announcements. In general, the market does not appear to react to CEO retirements: pre-retirement and post-retirement abnormal returns are not statistically different from zero. The result is consistent with the view that the market perceives CEO retirements (rather than CEO departures other than retirements) as orderly well-planned events and does not expect significant operating changes. The result is also in line with prior empirical research, as Huson et al. (2004) find no evidence of abnormal returns either before or after voluntary CEO departures.

To investigate whether the market reaction to CEO retirement is contingent on the structure of CEO SERP, I estimate abnormal returns around retirement announcement dates for the subsample of CEOs with performance-contingent SERPs. Pre-retirement abnormal returns are not significant, suggesting that the market does not differentiate between various types of CEO retirement arrangements prior to CEO retirement. As discussed earlier, a

TABLE 4
Long-Term Abnormal Stock Returns around CEO Retirement Announcements[a]

Panel A: Equally Weighted Portfolio[b]

	Pre-Retirement	**Post-Retirement**
All Retirements	0.00% (0.816)	−0.01% (0.734)
Implied Three-Year Abnormal Return	0.10%	−0.33%
Retirements with Performance-Contingent SERPs	0.06% (0.359)	−0.15% (0.093)
Implied Three-Year Abnormal Return	2.31%	−5.32%
Retirements with Performance-Contingent SERPs (*DACC* > median *DACC*)	0.05% (0.408)	−0.24% (0.035)
Implied Three-Year Abnormal Return	1.88%	−8.32%

Panel B: Value-Weighted Portfolio[b]

	Pre-Retirement	**Post-Retirement**
All Retirements	−0.03% (0.565)	−0.06% (0.380)
Implied Three-Year Abnormal Return	−1.07%	−2.22%
Retirements with Performance-Contingent SERPs	0.05% (0.401)	−0.11% (0.117)
Implied Three-Year Abnormal Return	1.83%	−3.92%
Retirements with Performance-Contingent SERPs (*DACC* > median *DACC*)	0.05% (0.394)	−0.22% (0.046)
Implied Three-Year Abnormal Return	1.82%	−7.59%

[a] This table reports monthly abnormal stock returns, estimated using calendar-time portfolio regressions, for the sample of CEOs of *Fortune* 1000 firms who retired between 1997 and 2006. p-values for two-tailed t-tests are reported in parentheses.

[b] Statistics for equally weighted (value-weighted) portfolios are presented in Panel A (Panel B). The pre-retirement (post-retirement) period includes 36 months prior to (after) the CEO retirement announcement. *DACC* for observations with performance-contingent SERPs refers to mean annual discretionary accruals during the pre-retirement period. The implied three-year abnormal return is calculated as $[(1 + \text{Monthly abnormal return})^{36} - 1]$.

---

[13] The reader is referred to Mitchell and Stafford (2000) for a detailed discussion of the calendar-time portfolio approach to measuring long-term abnormal performance.

potential reason is that the timing of CEO retirement may not be known to the market *ex ante*. In contrast, post-retirement abnormal returns are not only negative, consistent with the prediction, but also approach conventional levels of statistical significance. An average monthly abnormal return in the post-retirement period is -0.15 percent ($p < 0.10$) for the equally weighted portfolio and -0.11 percent ($p < 0.15$) for the value-weighted portfolio. Converting these monthly averages into implied three-year abnormal returns suggests that retirements of CEOs with performance-contingent SERPs lead to negative abnormal returns in the four to five percent range.

As further analysis, I partition the subsample of CEOs with performance-contingent SERPs into two equal subsamples based on average discretionary accruals in the pre-retirement period and investigate stock returns for the subsample with highest discretionary accruals. Results indicate that the market reaction to retirements of CEOs with performance-contingent SERPs depends on the size of discretionary accruals in the pre-retirement period. For observations with the highest pre-retirement accruals, an average monthly abnormal return in the post-retirement period ranges between -0.22 and -0.24 percent, leading to an implied three-year negative abnormal return of approximately eight percent. The results are statistically significant ($p < 0.05$) for both equally weighted and value-weighted portfolios.

In sum, I find no evidence of abnormal returns in the period preceding the retirement of CEOs with performance-contingent SERPs. This finding is consistent with the prediction that the market does not unravel income-increasing earnings management in these years, likely because of an *ex ante* information asymmetry between shareholders and the CEO regarding the timing of CEO retirement.[14] In contrast, in the period following the retirement of CEOs, performance-contingent SERPs are associated with negative long-term abnormal returns, with the significance and the magnitude of these returns positively related to the size of discretionary accruals in the pre-retirement period. Taken together, the results suggest that the equity of firms with performance-contingent CEO SERPs is temporarily overpriced in the period preceding CEO retirement.

V. CONCLUSIONS

Prior studies of the impact of managerial retirement on accounting discretion are inconclusive, with most empirical studies finding little or no support of the theoretical prediction of income-increasing earnings management in the pre-retirement period. I argue that an increase in accounting earnings in the pre-retirement period is particularly beneficial for managers whose Supplemental Executive Retirement Plan (SERP) benefits are contingent on firm performance in this period. An empirical analysis of CEO retirements from *Fortune* 1000 firms supports the hypothesized relationship between CEO retirement arrangements and firm discretionary accruals. Specifically, results provide evidence of income-increasing earnings management in the pre-retirement period in firms with performance-contingent

---

[14] Although the average number of determination years per CEO is close to 36 months (Table 1), the determination period may be as short as 12 months in individual cases. Consequently, it is possible that abnormal returns in the pre-retirement period are not detected because income-increasing earnings management does not exist in all 36 pre-retirement months in firms with performance-contingent CEO SERPs. To triangulate the findings, I investigate abnormal returns on pre-retirement portfolios constructed with a 12-month horizon (i.e., for each month t, all firms in which the CEO will retire within the next 12 months are included in that month's pre-retirement portfolio). The results show no evidence of abnormal returns in the 12-month period prior to CEO retirement regardless of the sample (all retirements, retirements with performance-contingent SERPs, or retirements with performance-contingent SERPs characterized by highest pre-retirement discretionary accruals) and portfolio construction (equally weighted or value-weighted).

CEO SERPs, but no evidence of earnings management associated with CEO retirement in other firms.

As further evidence of the consequences of earnings management in firms with performance-contingent CEO SERPs, I investigate the market reaction to CEO retirements. Consistent with prior studies, I find no evidence of abnormal returns in pre- or post-retirement periods when *all* CEO retirements are considered. However, a partitioned analysis indicates that the market reaction to CEO retirement depends on CEO retirement arrangements. Retirements of CEOs with no SERP and CEOs whose SERP benefits are not contingent on firm performance do not lead to abnormal market returns around CEO retirement announcements. In contrast, the post-retirement period in firms with performance-contingent CEO SERPs is characterized by negative long-term abnormal returns. The magnitude and the significance of these returns are positively related to the size of discretionary accruals in the pre-retirement period. Thus, results suggest that the market reacts to income-increasing earnings management in the pre-retirement period, but only after the CEO announces his/her retirement. This conclusion is consistent with the existence of an *ex ante* information asymmetry between shareholders and the CEO regarding the timing of CEO retirement. Findings also imply that the equity of firms with performance-contingent CEO SERPs is temporarily overpriced in the final years prior to CEO retirement.

This research makes several important contributions to the literature. First, it provides new empirical evidence on the impact of managerial retirement on earnings management practices. This contribution is particularly important in light of prior empirical studies that, in most cases, do not find significant relationships between managerial horizon and accounting discretion. Second, findings highlight the role of managerial compensation in corporate decisions and, more specifically, the impact of compensation contracts on earnings management incentives. Third, the results provide evidence of negative long-term abnormal returns in the post-retirement period in firms with performance-based CEO SERPs, thereby extending the research on the market reaction to earnings management and CEO retirements. Fourth, the study sheds light on the dynamics of the horizon problem, insofar as no prior study has demonstrated that the horizon problem arises as early as in the third year prior to CEO retirement and reaches its apex in the second year prior to CEO retirement. Finally, in the spirit of contingency theories, this research highlights the complexity of the phenomenon of accounting discretion and the importance of contextual factors in earnings management tests. In contrast to studies that investigate *either* the impact of managerial horizon *or* the impact of managerial compensation on discretionary choices, this research takes a more integrative approach to exploring earnings management incentives.

REFERENCES

Aboody, D., and R. Kasznik. 2000. CEO stock option awards and the timing of corporate voluntary disclosures. *Journal of Accounting and Economics* 29 (1): 73–100.

Barker, V., and G. Mueller. 2002. CEO characteristics and firm R&D spending. *Management Science* 48 (6): 782–801.

Barth, M., W. Beaver, J. Hand, and W. Landsman. 1999. Accruals, cash flows, and equity values. *Review of Accounting Studies* 4 (3–4): 205–229.

Bartov, E., and P. Mohanram. 2004. Private information, earnings manipulations, and executive stock-option exercises. *The Accounting Review* 79 (4): 889–920.

Becker, C., M. DeFond, J. Jiambalvo, and K. Subramanyam. 1998. The effect of audit quality on earnings management. *Contemporary Accounting Research* 15 (1): 1–24.

Butler, M., A. Leone, and M. Willenborg. 2004. An empirical analysis of auditor reporting and its association with abnormal accruals. *Journal of Accounting and Economics* 37 (2): 139–165.

Butler, S., and H. Newman. 1989. Agency control mechanism effectiveness and decision making in an executive's final year with a firm. *Journal of Institutional and Theoretical Economics* 145: 451–464.

Cheng, S. 2004. R&D expenditures and CEO compensation. *The Accounting Review* 79 (2): 305–328.

Chung, H., and S. Kallapur. 2003. Client importance, nonaudit services, and abnormal accruals. *The Accounting Review* 78 (4): 931–955.

Craighead, J., M. Magnan, and L. Thorne. 2004. The impact of mandated disclosure on performance-based CEO compensation. *Contemporary Accounting Research* 21 (2): 369–398.

DeAngelo, L. 1988. Managerial competition, information costs, and corporate governance, the use of accounting performance measures in proxy contests. *Journal of Accounting and Economics* 10 (1): 3–36.

Dechow, P., and R. Sloan. 1991. Executive incentives and the horizon problem: An empirical investigation. *Journal of Accounting and Economics* 14 (1): 51–89.

———, ———, and R. Sweeney. 1995. Detecting earnings management. *The Accounting Review* 70 (2): 193–225.

DeFond, M., and J. Jiambalvo. 1994. Debt covenant violation and manipulation of accruals. *Journal of Accounting and Economics* 17 (1–2): 145–176.

Denis, D., and D. Denis. 1995. Firm performance changes following top management dismissals. *The Journal of Finance* 50 (4): 1029–1057.

Fama, E., and K. French. 1993. Common risk factors in the returns on stocks and bonds. *Journal of Financial Economics* 33 (1): 3–56.

Frankel, R., M. Johnson, and K. Nelson. 2002. The relationship between audit fees for nonaudit services and earnings quality. *The Accounting Review* 77 (Supplement): 71–105.

Gaver, J., K. Gaver, and J. Austin. 1995. Additional evidence of bonus plans and income management. *Journal of Accounting and Economics* 19 (1): 3–28.

Gibbons, R., and K. Murphy. 1992. Does executive compensation affect investment? *Journal of Applied Corporate Finance* 5 (2): 99–109.

Guidry, F., A. Leone, and S. Rock. 1999. Earnings-based bonus plans and earnings management by business unit managers. *Journal of Accounting and Economics* 26 (1–3): 113–142.

Healy, P. 1985. The effect of bonus schemes on accounting decisions. *Journal of Accounting and Economics* 7 (1–3): 85–107.

Holthausen, R., D. Larcker, and R. Sloan. 1995. Annual bonus schemes and the manipulation of earnings. *Journal of Accounting and Economics* 19 (1): 29–74.

Hribar, P., and D. Collins. 2002. Errors in estimating accruals: Implications for empirical research. *Journal of Accounting Research* 40 (1): 105–134.

———, and D. Nichols. 2007. The use of unsigned earnings quality measures in tests of earnings management. *Journal of Accounting Research* 45 (5): 1017–1053.

Huson, M., P. Malatesta, and R. Parrino. 2004. Managerial succession and firm performance. *Journal of Financial Economics* 74 (2): 237–275.

Jensen, M., and K. Murphy. 1990. Performance pay and top-management incentives. *The Journal of Political Economy* 98 (2): 225–264.

Jones, J. 1991. Earnings management during import relief investigations. *Journal of Accounting Research* 29 (2): 193–228.

Kalyta, P., and M. Magnan. 2008. Executive pensions, disclosure quality, and rent extraction. *Journal of Accounting and Public Policy* 27 (2): 133–166.

———. 2009. Compensation transparency and managerial opportunism: A study of supplemental retirement plans. *Strategic Management Journal* 30 (4): 405–423.

Kim, J., R. Chung, and M. Firth. 2003. Auditor conservatism, asymmetric monitoring, and earnings management. *Contemporary Accounting Research* 20 (2): 441–464.

Klein, A. 2002. Audit committee, board of director characteristics, and earnings management. *Journal of Accounting and Economics* 33 (3): 375–400.

Kothari, S. P., A. Leone, and C. Wasley. 2005. Performance matched discretionary accrual measures. *Journal of Accounting and Economics* 39 (1): 163–197.

Lambert, R., D. Larcker, and K. Weigelt. 1991. How sensitive is executive compensation to organizational size? *Strategic Management Journal* 12 (5): 395–402.

McNichols, M., and P. Wilson. 1988. Evidence of earnings management from the provision for bad debts. *Journal of Accounting Research* 26 (supplement): 1–31.

Menon, K., and D. Williams. 2004. Former audit partners and abnormal accruals. *The Accounting Review* 79 (4): 1095–1118.

Mitchell, M., and E. Stafford. 2000. Managerial decisions and long-term stock price performance. *Journal of Business* 73 (3): 287–329.

Murphy, K., and J. Zimmerman 1993. Financial performance surrounding CEO turnover. *Journal of Accounting and Economics* 16 (1–3): 273–315.

———. 2000. Performance standards in incentive contracts. *Journal of Accounting and Economics* 30 (3): 245–278.

Pourciau, S. 1993. Earnings management and nonroutine executive changes. *Journal of Accounting and Economics* 16 (1–3): 317–336.

Smith, C., and R. Watts. 1982. Incentive and tax effects of executive compensation plans. *Australian Journal of Management* 7 (2): 139–157.

Sundaram, R., and D. Yermack. 2007. Pay me later: Inside debt and its role in managerial compensation. *The Journal of Finance* 62 (4): 1551–1588.

Vancil, R. 1987. *Passing the Baton: Managing the Process of CEO Succession.* Boston, MA: Harvard Business School Press.

Watts, R., and J. Zimmerman, eds. 1986. *Positive Accounting Theory*. Englewood Cliffs, NJ: Prentice Hall.

Weisbach, M. 1988. Outside directors and CEO turnover. *Journal of Financial Economics* 20 (1–2): 431–460.

Wells, P. 2002. Earnings management surrounding CEO changes. *Accounting and Finance* 42 (2): 169–193.

Yermack, D. 1997. Good timing: CEO stock option awards and company news announcements. *The Journal of Finance* 52 (2): 449–476.

THE ACCOUNTING REVIEW
Vol. 84, No. 5
2009
pp. 1575–1606

American Accounting Association
DOI: 10.2308/accr.2009.84.5.1575

Management Forecasts in Japan: An Empirical Study of Forecasts that Are Effectively Mandated

Kazuo Kato
Osaka University of Economics

Douglas J. Skinner
The University of Chicago

Michio Kunimura
Meijo University

ABSTRACT: We study management forecasts in Japan, where forecasts are effectively mandated but managers have considerable latitude over the numbers they release. We find that managers' initial earnings forecasts for a fiscal year are systematically upward-biased but that they revise their forecasts downward during the fiscal year so that most earnings surprises are non-negative. Managers' initial forecast optimism is inversely related to firm performance, and is more pronounced for firms with higher levels of insider ownership, smaller firms, and firms with a history of forecast optimism. The fact that managers' forecasts tend to be consistently optimistic suggests that reputation effects are insufficient to ensure managerial forecast accuracy. We also find that the information content of managers' forecasts is related to proxies for whether market participants view the forecasts as credible.

Keywords: *management forecasts; Japan; forecast credibility; forecast optimism.*

Data Availability: *The data used in this study are publicly available from the sources indicated herein.*

We are grateful to two anonymous referees and workshop participants at the Asian Academic Accounting Association Conference (Sydney), Australasian Finance and Banking Conference, Baruch College–CUNY, University of Chicago, Macquarie University, Michigan State University, University of New South Wales, Stanford University, University of Sydney, and The University of Texas at Dallas for helpful comments. Kazuo Kato acknowledges financial support from the Japanese Society for the Promotion of Science, Granted-in-Aid for Scientific Research C16530310 and C19530421. This article was previously titled "When Voluntary Disclosure Isn't Voluntary: Management Forecasts in Japan."

Editor's note: Accepted by Steven Kachelmeier, with thanks to Dan Dhaliwal for serving as editor on a previous version.

Submitted: April 2007
Accepted: December 2008
Published Online: September 2009

I. INTRODUCTION

We study management forecasting in Japan. This setting allows us to investigate the credibility of management forecasts in an economy where managers are required to provide forecasts but have considerable latitude over the numbers they release. While the credibility of management forecasts has long been a central issue in the disclosure literature (King et al. 1990), most research has been conducted in settings where forecasts are voluntary and there are clear legal sanctions for providing false or misleading disclosures and omissions. In Japan, managers are required to provide initial forecasts at the beginning of the fiscal year and then to revise those forecasts during the period. Because the regulatory and legal costs of biasing forecasts are relatively low in Japan, managers can bias their initial forecasts but then revise their forecasts later in the fiscal period to avoid earnings surprises.

In contrast to research based on U.S. data (e.g., McNichols 1989), we find that, on average, Japanese managers bias their initial forecasts upward but then revise their forecasts downward during the fiscal period so that most earnings surprises are non-negative. We show that this initial forecast optimism is inversely related to firm performance—managers of poorly performing firms provide the most optimistic forecasts. We then investigate how this bias affects forecast credibility by looking at the stock price effects associated with these forecasts. We find that, on average, initial management forecasts are informative but that forecast informativeness is lower for firms with poor performance, consistent with them being less credible.

Because forecasts in Japan are effectively mandated, we have a panel of firm-level forecasts, and so can investigate how managers' forecast bias affects the credibility of their subsequent forecasts. There is a strong tendency for managers whose forecasts are overly optimistic in one year to be too optimistic again in subsequent years. We find that this optimism reduces stock market participants' response to the managers' subsequent forecasts, consistent with the importance of reputation effects.

We discuss a number of explanations for managerial forecast optimism, including overconfidence (Malmendier and Tate 2005), behavioral biases (Kahneman and Lovallo 1993), the possibility that forecasts are tied to internal budgets used to motivate employees, and managerial opportunism (poorly performing managers portray their firms' performance as overly favorable). Our results are most consistent with opportunism. We find that optimism is more pronounced for firms with poor past performance, for firms with relatively high levels of managerial/insider ownership, and is less pronounced for larger firms. These variables likely capture cross-sectional variation in the extent of managerial monitoring. On the other hand, the persistent nature of forecast optimism makes the results less consistent with opportunism because it seems less likely that those outside the firm will be "fooled" by persistent optimism.

Overall, our findings suggest that in an environment where managers' forecasts are both effectively mandated and relatively unconstrained by legal or regulatory forces, managers tend to provide forecasts that are biased upward and that this optimism is persistent. We also find that while these forecasts are, on average, informative to investors, investors discount forecasts issued by managers whose firms are performing poorly or who have issued overly optimistic forecasts in the past, suggesting that past over-optimism reduces forecast credibility.

Our results contribute to the literature on management forecasts in three ways. First, there is little or no previous evidence on the properties of managers' earnings forecasts in a setting where forecasts are mandated rather than voluntary. Second, there is little extant evidence on whether market forces such as reputation are sufficient to discipline managers' forecasts in the absence of legal and/or regulatory forces. Third, there is limited evidence on whether factors that are likely to affect the credibility of management forecasts (such as firm performance and managerial reputation) are priced by market participants. By studying managerial forecasts in Japan, we provide evidence on each of these issues.

Section II lays out the Japanese institutional setting in more detail and develops our hypotheses. Section III describes our sample. Section IV reports our main empirical analyses. Section V concludes.

II. INSTITUTIONAL BACKGROUND AND HYPOTHESIS DEVELOPMENT

Management Forecasts in Japan

The Stock Exchange Act (the Act) in Japan governs disclosure and financial reporting practices for Japanese public companies. In addition, the so-called Timely Disclosure Rules (*Kessan-Tannsin* or "summary of financial statements") enforced by Japanese stock exchanges impose more stringent requirements on disclosure practices. These rules originated in 1965 from the "*Kabuto*-club," a club of newspapermen at the Tokyo Stock Exchange (TSE), and were eventually incorporated into TSE rules.[1] These rules strongly encourage managers of listed firms in Japan to provide regular forecasts of sales and earnings. Our evidence indicates that the large majority of companies comply with this request and so we argue that forecast disclosure is effectively mandated in Japan. The specific requirements are as follows:[2]

(1) Listed companies are expected to release point forecasts of annual earnings at each annual earnings announcement date, as well as revisions of these forecasts at interim earnings announcement dates. Thus, managers provide initial forecasts for year t when year $t-1$ earnings are announced, and revisions (which include confirmations) when interim earnings are announced.

(2) Managers are expected to provide forecasts of Sales, Earnings before Extraordinary items and Taxes (EBET), and Net Income (NI).

(3) Forecasts must be updated if there are "significant" revisions in management estimates, defined as changes in sales estimates of 10 percent or more and/or changes in earnings estimates of 30 percent or more (the "Significance Rule"). In contrast to initial forecasts, which are encouraged by stock exchange listing rules, these revisions are required under the Act.

Hypotheses

The credibility of managers' earnings forecasts has been a central issue in the disclosure literature for some time (King et al. 1990). For example, Pownall and Waymire (1989) compare the credibility of management forecasts (which are voluntary) to that of earnings announcements (which are mandated). Because earnings realizations are likely to be more reliable than earnings forecasts, Pownall and Waymire (1989) expect that earnings announcements are likely to be more credible than management forecasts. These authors find, however, that the stock price response associated with earnings forecasts is stronger than that associated with earnings announcements, a result they attribute to greater forecast timeliness.[3]

---

[1] In their original form, these rules required forecasts of sales, net income, and dividends, along with a summary of financial statement information. The rules are similar on all of the major Japanese stock exchanges, including the TSE, Osaka Stock Exchange, and the JASDAQ.

[2] Several previous papers also look at management forecasts in Japan. Darrough and Harris (1991) use management forecast data from 1979–1987 to examine the information content of Japanese management forecasts. Conroy et al. (1998) use data from 1983–1995 to examine hypotheses about the extent to which Japanese stock prices are driven by "fundamentals" (earnings) while Conroy et al. (2000) use largely the same data to assess the marginal information content of dividends over earnings.

The management forecast literature has largely developed in the U.S., where management earnings forecasts are voluntary. Because forecasts are voluntary, the literature devotes a great deal of attention to managers' incentives to provide forecasts. Ajinkya and Gift (1984) put forward the "expectations adjustment hypothesis" under which managers provide forecasts when their expectations diverge from those of market participants such as analysts. In addition, theory suggests that managers of firms with "good news" (above a certain threshold) disclose that news, whereas managers of other firms withhold earnings news (Verrecchia 1983). The evidence supports this view, especially in earlier time periods (Patell 1976; Penman 1980; Lev and Penman 1990). There is also evidence that managers with adverse earnings news issue forecasts to preempt the market reaction to that news, for legal and/or reputational reasons (Skinner 1994; Kasznik and Lev 1995).

These results suggest that managers tend to issue earnings forecasts when their earnings expectations are different from those of outside investors. Because market participants, on average, are likely to have rational expectations, the market's reaction to earnings forecasts is likely to depend on what outsiders infer from the decision to forecast about managers' underlying earnings news.

This endogeneity makes it hard to use evidence from voluntary disclosures to make inferences about the credibility of management forecasts in general. Because previous research generally shows that managers' incentives to provide forecasts increase with the magnitude of the earnings news, it is not surprising that these disclosures are, on average, informative. But we cannot infer from this that managers' forecasts are always informative. Because forecasts in Japan are effectively mandated, there is no information in the forecast decision itself.

We first investigate how managers set their initial earnings forecasts in each fiscal year. While there is a great deal of evidence on the properties of voluntary forecasts, we know little about how managers arrive at their earnings forecasts, especially those made early in the fiscal period before any news about current period performance is available.

Management can take one of three possible approaches to setting forecasts: (1) managers can set forecasts in an unbiased manner, based on their best estimates of earnings for the period, (2) managers can make downward-biased (conservative) forecasts, or (3) managers make upward-biased (optimistic) forecasts. Which of these three possibilities is most descriptive depends on the benefits and costs to managers of biasing earnings forecasts.

Because of the nature of the legal and regulatory environment in Japan, the legal costs of biasing earnings forecasts are likely to be relatively small.[4] There may, however, be reputation costs. Managers who issue forecasts that are consistently biased are likely to lose credibility with market participants, which reduces the effectiveness of their future communications and possibly also their overall reputation as managers (Hutton and Stocken 2006).

---

[3] Atiase et al. (2005) investigate a similar question using concurrent announcements of earnings and earnings forecasts. They find that earnings news is more highly associated with stock returns.

[4] Litigation generally in Japan, including securities litigation, has traditionally been much less prevalent than in Western countries (Ginsburg and Hoetker 2006). In spite of an increase in litigation rates since around 1990, it is still the case that expected litigation costs are lower in Japan than in the U.S.; for example, West (2001) provides evidence that the number of shareholder derivative lawsuits has increased but that settlements are unusual and that stockholders lose most of these cases. In addition, Japan has not had, until recently, a securities regulator analogous to the U.S. SEC. Until 2000, the enforcement of securities laws was handled by the Ministry of Finance (MOF), which was also responsible for promoting business and managing the economy. In 2000, the Financial Services Agency (FSA) was created to handle enforcement of securities and banking laws. There is little evidence, however, that it enforces securities laws, including disclosure requirements, nearly as rigorously as the U.S. SEC. See for example Markham (2003).

Japanese managers will have incentives to provide conservative initial forecasts if negative forecast revisions and/or earnings surprises are more costly than positive revisions/surprises. Skinner (1994) provides evidence that managers in the U.S. disclose earnings news in a way that suggests they face an asymmetric loss function, and attributes this to legal and/or reputational costs. Because Japanese managers do not face legal costs from adverse earnings surprises, a tendency for them to issue conservative forecasts would likely be due to reputational forces.

It seems more likely that Japanese managers make upward-biased (optimistic) forecasts. First, if viewed as credible by capital market participants, optimistic forecasts increase stock prices, at least in the short term. Managers' incentives to make disclosures that increase stock price receive prominent attention in the literature (Verrecchia 1983). Second, providing good news about the firm's future earnings prospects helps managers convince the firm's constituents (stockholders, the board, lenders, employees, suppliers, etc.) that they are doing a good job.[5] If this is the case, then the benefits of overly optimistic forecasts are likely to be greater for managers of firms whose current performance is poor.[6] Consistent with this, Rogers and Stocken (2005) provide U.S. evidence that managers of distressed firms are more likely to issue optimistic forecasts than managers of other firms.

Managers' incentives to bias forecasts are also likely to depend on whether their firms are part of "main bank" or *keiretsu* groups.[7] In these cases the holdings of group banks, insurance companies, and other large group stockholders give them incentives to monitor managers more closely than if ownership is dispersed (Shleifer and Vishny 1986). Alternatively, managers of firms with large group cross-holdings may be subject to *less* external discipline if these stockholders have access to internal company information, especially if they are affiliated with management (Fama and Jensen 1983; Morck et al. 1988).[8] To investigate these predictions, we provide evidence on whether forecast bias varies cross-sectionally with measures of shareholder ownership.

An alternative perspective is that managers' earnings forecasts are attributable to the firm's internal budgeting/performance evaluation system and that their use as external forecasts is merely a by-product of this process. Under this view, we also expect forecasts to be optimistic, in the sense of being at least as large as current realizations.[9] In addition, internal budget targets may be set optimistically to motivate the firm's managers and employees.[10] This argument implies that forecasts will be relatively more optimistic for firms with poor performance, and that managerial optimism may be persistent even if this makes forecasts less credible to outsiders.

---

[5] In Japan, initial forecasts are issued soon before the annual stockholders' meeting, when managers are likely be especially concerned with how firm performance is perceived.

[6] Kaplan (1994) finds that management turnover in Japan increases with poor earnings and stock price performance and is especially sensitive to low reported earnings.

[7] For background on the "main bank" system in Japan see Aoki et al. (1994) and Hoshi and Kashyap (2001).

[8] It is often the case that executives of the main bank and other group companies sit on the board of the companies in which they hold shares. In addition, because main banks also provide debt financing to group companies, they likely receive internal information in their role as lenders.

[9] Under budget "ratcheting," targets are based on actual performance when it exceeds the past target and on the past target when there is a shortfall. This means that forecasts will be at least as large as current realizations, and larger than this if past performance is poor. See Leone and Rock (2002) or Zimmerman (2006, Chap. 6).

[10] These arguments may be particularly relevant in Japan given the prevalence of *kaizen* (continuous improvement) costing systems. Under *kaizen* costing, budgets are set based on continuous improvements in production efficiency, so that budget targets for one period are unlikely ever to be set below actual performance in the prior period, and are often well above these levels. See, for example, Cooper (1995, Chap. 11) and Cooper and Kaplan (1998, Chap. 5).

Another possibility is that some form of cognitive bias causes managers to be unduly optimistic in making their initial forecasts. Managers of firms that are doing well may rely too heavily on their own firm's past performance to make forecasts, and underweight the tendency for unusual performance to revert to the mean (Lakonishok et al. 1994). This form of bias predicts that managers of firms that are doing well overweight the possibility that performance will persist, and *vice versa*. Alternatively, it may be that managers, in setting their forecasts, place too much weight on specific internal idiosyncratic information about firm performance and underweight external information (Kahneman and Lovallo 1993). This form of bias could result in a negative relation between past performance and forecast optimism if managers erroneously over-weight internal forecasts that imply an improvement in performance. Finally, forecast optimism could be due to managerial over-confidence, which has recently been linked to a number of corporate decisions (Malmendier and Tate 2005; Ben-David et al. 2007).

These arguments lead to our first three hypotheses, stated in null form:

H1: Managers' initial earnings forecasts are unbiased.

H2: Other factors held constant, there is no relation between the bias in managers' initial earnings forecasts and their firms' contemporaneous realized earnings performance.

H3: Other factors held constant, there is no relation between the ownership structure of Japanese firms and their managers' incentives to provide biased initial forecasts.

We argue above that Japanese managers have considerable latitude in setting their initial forecasts because they can subsequently revise their forecasts at low cost and because the legal/regulatory costs of issuing biased forecasts are relatively low. This raises the question of whether these forecasts are credible to those outside the firm.[11] We therefore provide evidence on whether managers' initial forecasts are informative to market participants, and whether the informativeness of these forecasts depends on their incentives to bias those forecasts. In null form:

H4: Managers' initial earnings forecasts are uninformative to market participants.

H5: The extent to which managers' forecasts are informative is unrelated to their incentives to bias those forecasts.

Although the empirical setting is different, H5 is similar to the hypothesis that Rogers and Stocken (2005) investigate because it implies that investors condition their reaction to forecasts on managers' incentives to bias the forecasts. To the extent that we find support for H2—that managers' forecast bias is negatively related to firm performance—we expect that past performance influences forecast credibility.

We do not collect analysts' forecasts for these firms because our main hypotheses are about how managers set forecasts at the beginning of the fiscal year, which we believe is most likely affected by contemporaneous firm performance and past forecast accuracy, rather than market expectations. Moreover, the fact that we do not know a great deal about analysts' forecasts in Japan (whether analysts have incentives to be accurate, whether they cater to management, etc.) makes it difficult to develop clear predictions about the interaction between management forecasts and analysts' forecasts. We acknowledge that this is a potential limitation of our study; to the extent that managers' forecasts are affected by analysts' forecasts or by how analysts are likely to respond to managers' forecasts, we may miss important determinants and/or consequences of management forecasts.

---

[11] Stocken (2000) shows that even when managers' disclosures are cheap talk, in a multi-period setting reputational considerations may be sufficient to induce managers to provide truthful (unbiased forecasts).

Our next hypothesis concerns potential autocorrelation of managers' forecast bias. An advantage of the Japanese setting is that we have time-series data on managers' forecasting decisions. If reputation is an important consideration to managers, we expect that managers who make optimistic forecasts in one year are less likely to issue optimistic forecasts in the next year (Stocken 2000). We test the following null hypothesis:

> **H6:** Other things held constant, there is no relation between the bias in managers' initial earnings forecasts for year t and the bias in their initial forecasts in year $t+1$.

On the other hand, if forecasts are largely the result of internal budget-setting processes, then there is no reason to expect that over-optimism in one period will reverse in a subsequent period. The same argument likely applies to managerial behavioral biases or overconfidence; i.e., it is unclear why bias due to these factors would reverse over time.[12]

The discussion to this point relates to managers' initial fiscal-year forecasts. We make two predictions regarding forecast revisions. First, to the extent that managers' initial forecasts are biased and earnings surprises are costly, we expect that managers revise their forecasts toward realizations so that earnings surprises are relatively small. This result is also expected because managers obtain more information about actual performance as the year progresses, which naturally makes their forecasts more accurate. Because the alternative to this argument (that managers' forecasts become less accurate as the year progresses) is implausible, we do not test it as a formal hypothesis.

Second, to the extent that managers' initial forecasts are biased and market participants discount the forecast news, we expect forecast revisions that correct the initial bias are more informative than the initial forecasts, and that this difference increases with the extent to which the initial forecast is biased. This results in the following null hypothesis:

> **H7:** Managers' initial forecasts and subsequent revisions are equally informative.

III. SAMPLE AND DESCRIPTIVE INFORMATION

Our data are from Nikkei Financial Quest, a commercial database provided by Nikkei, publisher of the main business newspaper in Japan. The sample of firm/years with forecast data is shown in Panel A of Table 1 and extends from 1997 to 2007 (fiscal years). To include a given firm/year t in the sample, we require forecasts for year t and realizations for year $t-1$ (we also have realized year t earnings for all years except fiscal 2007). The sample covers most listed firms in Japan. The second column in the table shows the number of firm/years in the database with available data. The third column shows the number of these firm/years for which we also have forecast data. These numbers show that over 90

---

[12] It could be that managers become more competent forecasters over time, so that forecast bias due to overconfidence or behavioral biases is reduced.

TABLE 1
Sample and Descriptive Statistics

Panel A: Initial Management Forecasts[a]

Year	All Available Firm-Years	Firm-Years with Management Forecasts	Percentage of Available Firm-Years with Forecast
1997	3,240	2,832	87.4
1998	3,313	3,159	95.4
1999	3,367	3,226	95.8
2000	3,488	3,297	94.5
2001	3,599	3,402	94.5
2002	3,731	3,520	94.3
2003	3,798	3,574	94.1
2004	3,889	3,619	93.1
2005	4,010	3,728	93.0
2006	4,119	3,824	92.8
2007	4,093	3,887	95.0
Total	40,647	38,068	93.7

(continued on next page)

TABLE 1 (continued)

Panel B: Descriptive Statistics for Firms Included in Sample of Initial Management Forecasts[b]

	1997	1998	1999	2000	2001	2002	2003	2004	2005	2006
Obs.	2,875	2,748	2,947	3,058	3,114	3,197	3,285	3,361	3,455	3,654
Total Assets	200,008	201,441	342,423	352,554	401,012	396,756	399,507	384,580	427,127	419,689
(¥ million)	(34,256)	(38,113)	(37,217)	(32,753)	(30,287)	(30,048)	(29,767)	(30,355)	(29,649)	(29,649)
Debt/Equity Ratio	5.295	3.847	2.926	3.620	3.198	3.015	2.600	2.218	1.887	2.232
	(1.480)	(1.438)	(1.410)	(1.443)	(1.352)	(1.312)	(1.273)	(1.205)	(1.136)	(1.132)
ROA	0.010	0.001	0.009	0.008	−0.001	0.006	0.015	0.021	0.023	0.019
	(0.012)	(0.008)	(0.012)	(0.012)	(0.008)	(0.013)	(0.020)	(0.025)	(0.027)	(0.028)
Market Cap	86,667	96,959	133,375	101,810	86,241	65,353	91,873	94,603	141,300	137,013
(¥ million)	(12,300)	(11,187)	(11,671)	(10,555)	(9,185)	(7,810)	(12,600)	(16,083)	(20,736)	(15,805)

---

[a] Panel A shows the number of firm-years in the sample of initial management earnings forecasts (the initial forecast being the first forecast for a given fiscal period). Each observation is classified by fiscal year, which usually ends on March 31 for Japanese companies; 1997 thus denotes the year ended March 31, 1998. To be included in the sample, the firm must be listed on one of the following stock exchanges at the end of a given fiscal year: Tokyo, Osaka, Nagoya, and/or regional exchanges including JASDAQ (except for 2007, for which we look at the beginning of the year). The sample includes financial institutions. "All Available Firm-Years" are firm-years for which realized earnings data are available. "Firm-years with Management Forecasts" denotes the subset of these firm-years for which we also have data on management forecasts.

[b] Panel B reports sample means with sample medians in parentheses. All variables are measured at the end of each fiscal year and are from *Nikkei Financial Quest.*

Variable Definitions:

Debt/Equity = ratio of total liabilities to stockholders' equity; and
ROA = ratio of net income to lagged total assets; and
Market Cap = total market capitalization measured as shares outstanding multiplied by end of year closing stock price.

percent of firm/years have forecast data, which supports our contention that forecasting is effectively mandated in Japan.[13]

Panel B of Table 1 reports descriptive statistics for the sample, presented by year. Mean (median) total assets is ¥200 billion (¥34 billion) at the beginning of the period and ¥420 billion (¥30 billion) at the end of the period, while mean (median) market capitalization increases from $87 billion (¥12 billion) to $137 billion (¥16 billion). Several of the reported series reflect the weak state of the Japanese economy during the 1990s and the improvement that began after 2002. The overall mean (median) debt-to-equity ratio for these firms is 5.30 (1.48) in F1997 and declines to 2.23 (1.13) in F2006. Mean (median) profitability (ROA) is 0.1 percent (0.8 percent) in 1998 compared to 1.9 percent (2.8 percent) in 2006.

The data in Table 1 describe firm/years for which we have at least the initial annual forecast and past realized earnings. We use these data to test H1–H5. Beginning in fiscal 2002 Nikkei began collecting forecast revisions as well, which we use to provide evidence on how managers revise their forecasts during the fiscal period. We describe forecast revisions below.

IV. HYPOTHESIS TESTS

Forecast Optimism and its Determinants

To provide evidence on whether managers' initial forecasts are biased (H1), Table 2 reports forecast innovations and forecast errors. Forecast innovations are forecast earnings for year t minus realized earnings for year $t-1$, deflated by lagged total assets. Forecast errors are realized earnings for year t minus forecast earnings for year t, also deflated by lagged total assets. Throughout the paper we provide results for Sales and NI forecasts; results for EBET are generally similar to those for NI. To investigate whether managers' forecast bias varies as a function of contemporaneous realized performance (H2), we partition the data into quintiles based on realized profitability in year $t-1$, where Quintile 1 contains firm/years with the lowest profitability, as measured by return on assets (ROA).

The results in Table 2 show that managers' Sales and NI forecasts are optimistic. On average, forecast innovations are relatively large and positive, while forecast errors are relatively large and negative. We discuss forecast innovations first. For the sample overall, Sales are forecast to increase by a mean (median) 6.4 percent (3.3 percent) of assets, while NI is forecast to increase by 1.7 percent (0.5 percent) of assets. For both variables, over three-quarters (79.0 percent and 77.3 percent, respectively) of innovations are positive. These results are not especially surprising because we expect that, on average, Sales and NI increase.

What is more interesting in light of H2 is the variation across the ROA deciles and how this variation differs between Sales and NI forecasts. We find that managers of the most profitable firms (Quintile 5) forecast the largest Sales increases (mean 11.5 percent, median 6.2 percent) and that the magnitude of this increase declines with profitability. For NI, however, managers of the *least* profitable firms (Quintile 1) forecast the largest increases in NI (mean 7.4 percent, median 3.9 percent) while managers of the most profitable firms forecast essentially flat earnings (mean −0.1 percent, median 0.4 percent). Moreover, managers in the lowest ROA quintile forecast that NI will increase by more than Sales (mean 7.4 percent, median 3.9 percent for NI versus 5.0 percent and 2.6 percent for Sales). In all of the other quintiles the forecast increases in NI are substantially smaller than those for

---

[13] Missing observations arise either because the firm does not provide a forecast or because no forecast is available on the database. A few industries (such as the securities industry) are exempt from forecast requirements.

TABLE 2
Forecast Innovations and Forecast Errors for Initial Management Earnings Forecasts[a]

Panel A: Sales

	Overall		**Quintile 1 (lowest ROA_{t-1})**		**Quintile 2**		**Quintile 3**		**Quintile 4**		**Quintile 5 (highest ROA_{t-1})**	
Innovations[b] for Initial Management Forecasts												
Mean	0.064*		0.050*		0.042*		0.047*		0.065*		0.115*	
Median	0.033*		0.026*		0.023*		0.027*		0.036*		0.062*	
Obs.	33,843		6,771		6,770		6,767		6,770		6,765	
Positive	26,724	79.0%	4,829	71.3%	5,078	75.0%	5,281	78.0%	5,618	83.0%	5,918	87.5%
Forecast Errors[c] for Initial Management Forecasts												
Mean	−0.024*		−0.049*		−0.027*		−0.015*		−0.015*		−0.015*	
Median	−0.016*		−0.032*		−0.018*		−0.011*		−0.010*		−0.011*	
Obs.	30,644		6,131		6,130		6,128		6,130		6,125	
Positive	11,849	38.7%	1,873	30.5%	2,216	36.2%	2,517	41.1%	2,572	42.0%	2,671	43.6%

(*continued on next page*)

TABLE 2 (continued)

Panel B: Net Income

	Overall		Quintile 1 (lowest ROA_{t-1})		Quintile 2		Quintile 3		Quintile 4		Quintile 5 (highest ROA_{t-1})	
Innovation[b] for Initial Management Forecasts												
Mean	0.017*		0.074*		0.006*		0.003*		0.002*		−0.001	
Median	0.005*		0.039*		0.004*		0.002*		0.002*		0.004*	
Obs.	33,802		6,765		6,760		6,760		6,760		6,757	
Positive	26,140	77.3%	6,428	95.0%	5,514	81.6%	4,921	72.8%	4,657	68.9%	4,620	68.4%
Forecast Errors[c] for Initial Management Forecasts												
Mean	−0.014*		−0.036*		−0.013*		−0.008*		−0.007*		−0.008*	
Median	−0.003*		−0.009*		−0.003*		−0.002*		−0.002*		−0.002*	
Obs.	30,605		6,125		6,121		6,121		6,121		6,117	
Positive	11,395	37.2%	1,723	28.1%	2,077	33.9%	2,337	38.2%	2,549	41.6%	2,709	44.3%

---

* Denotes statistical significance at 1 percent level, two-tailed tests.

[a] The table reports sample means, medians, number of observations, and number and percentage of positive observations. Sample firm/years have year-ends between January 1998 and March 2008 for innovations and year-end between January 1998 and March 2007 for forecast errors.

[b] Innovation is defined as [management forecast of year t earnings − realized earnings for year $t-1$]/[total assets at $t-1$ year-end].

[c] Forecast Error is defined as [realized earnings for year t − management forecast of year t earnings]/[total assets at $t-1$ year-end].
Sample firm/years are divided into quintiles based on the magnitude of ROA for year $t-1$.

Sales. This suggests that managers of the poorest performers issue the most aggressive NI forecasts.[14]

One explanation for these results is mean reversion in earnings—it is well known that large earnings changes, decreases in particular, are transitory.[15] This is unlikely to fully explain these results, however, because they also hold for EBET, which excludes most unusual gains and losses (not tabulated) although to a lesser extent, and because of the monotonic relation across quintiles—while mean reversion may explain the extremes, it less easily explains the trend across quintiles.[16]

While the forecast innovations show that most Japanese managers forecast improvements in sales and earnings, the forecast errors show that most of these forecasts fall short of subsequent realizations. For Sales, the overall mean (median) forecast error is −2.4 percent (−1.6 percent), with only 38.7 percent positive. The largest (most negative) forecast errors are evident for firms in Quintile 1 (mean −4.9 percent, median −3.2 percent) with a close to monotonic increase through Quintile 5 (mean −1.5 percent, median −1.1 percent). There is also a monotonic decrease in the fraction of negative forecast errors as we move from Quintile 1 to Quintile 5 (from 69.5 percent to 56.4 percent), although the majority of forecast errors are negative across the quintiles and for the sample as a whole. These results show that managers are generally too optimistic about Sales, and that this optimism is inversely related to current profitability, consistent with H2.

Managers' NI forecasts are also upward biased. The mean (median) forecast error is −1.4 percent (−0.3 percent) and 63.8 percent of forecast errors are negative. Managers of the poorest performers are again the most optimistic. For Quintile 1 firms, the mean (median) forecast error is −3.6 percent (−0.9 percent) and 71.9 percent of the forecast errors are negative. Similar to Sales, over-optimism declines with profitability, although the majority of forecast errors are negative for all groups. The clear impression is one of upward-biased NI forecasts, especially for managers of poor performers.

Rather than reflecting managerial optimism, it is possible that the negative forecast errors are attributable to years when the Japanese economy unexpectedly changed for the worse after initial forecasts were issued. To assess this, Figure 1 plots the fraction of positive forecast innovations and forecast errors for each year for Sales and NI. The figure shows that forecast innovations for both variables are consistently positive throughout the sample period, with the fraction of positive observations varying between approximately 70 percent and 80 percent. There is greater year-to-year variation in forecast errors—the fraction of positive forecast errors increases from well below 50 percent in the years before 2003 to close to 50 percent for 2003 through 2006. This indicates that the systematic optimism evident in Table 2 is stronger in the first part of the sample period when there was an economic downturn in Japan.

When we replicate Table 2 using only observations from 2002–2007 (not reported in tables), the forecast errors are still negative, but less so than for the full period.[17] The trend across ROA quintiles is still strong, however. For example, the fraction of negative forecast

---

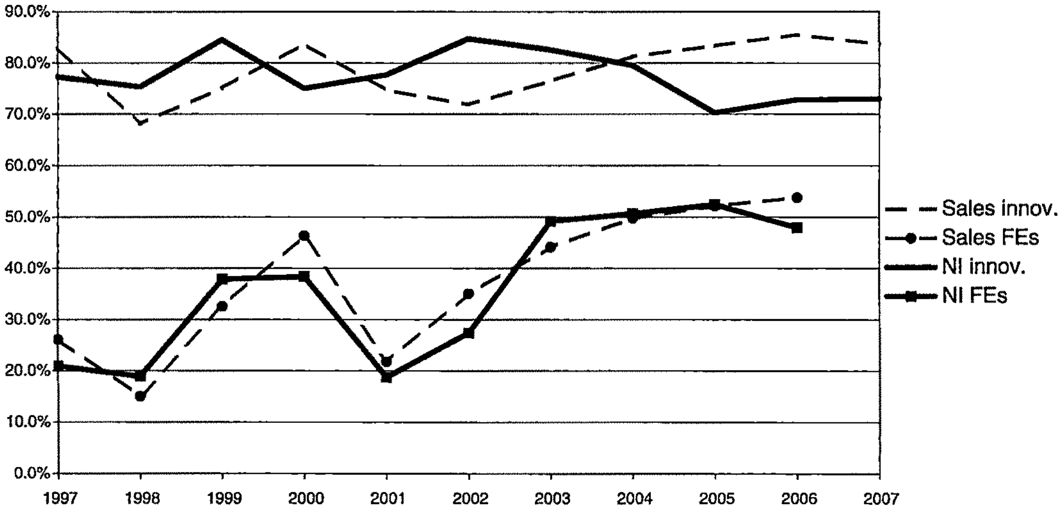
[14] These trends are also borne out by the proportion of positive observations in each cell: while the proportion of positive Sales innovations increases across the quintiles, from 71.3 percent in Quintile 1 to 87.5 percent in Quintile 5, exactly the opposite pattern holds for NI innovations, for which the fraction positive declines from 95.0 percent in Quintile 1 to 68.4 percent in Quintile 5.

[15] See, for example, Brooks and Buckmaster (1976), Freeman et al. (1982), and Fama and French (2000).

[16] To further address mean reversion, we re-performed the analyses in Table 2 after removing all observations for which year t–1 earnings are negative (much of mean reversion is due to loss reversals). Our results are unchanged.

[17] The overall mean (median) forecast error for NI during this period is –1.3 percent (–0.1 percent), with 54.5 percent of the errors negative; these numbers are significantly negative at the 1 percent level.

FIGURE 1
Fraction of Managers' Initial Forecast Innovations and Forecast Errors that Are Positive



errors for NI is 65 percent in Quintile 1 and declines monotonically across quintiles. Thus, while managers' overall forecast optimism is largely attributable to the poor economic conditions in Japan during the 1990s, this cannot explain the strong cross-sectional relation between performance and forecast optimism. We account for the systematic time-series variation in the regression analyses that we report below.

Forecast Credibility and Reputation Effects

We next provide evidence on managers' forecast strategies over time. We expect that managers whose initial forecasts are overly optimistic in year $t-1$ are less likely to be overly optimistic in year t because of the adverse effect persistent over-optimism is likely to have on their reputation and forecast credibility. To address this, Table 3 presents two 2 × 2 tables that report the number of observations (and percentage of these observations that are positive, in parentheses) of year t forecast innovations (Panel A) and year t forecast errors (Panel B) in each of four categories obtained by sorting the data according to the sign of year $t-1$ forecast innovations and forecast errors. Thus, for example, Panel A shows that of the set of 8,299 observations for which year $t-1$ forecast errors and forecast innovations are both non-negative (NW cell), 63.2 percent of year t forecasts represent positive innovations. The table is based on NI forecasts; results for EBET and Sales are similar.

The data in Table 3 strongly suggest that managers' forecast optimism is persistent. Panel A shows that when managers' year $t-1$ forecast innovations are non-negative *and* result in negative forecast errors (SW cell: the over-optimism case), managers are very likely to again forecast a positive innovation in year t (87.5 percent of observations in this cell are positive), and this forecast is again likely to be too optimistic (in 68.6 percent of cases; SW cell of Panel B). Moreover, regardless of the sign of the year $t-1$ innovation, managers whose forecasts result in negative forecast errors in year $t-1$ (total of 18,022 observations) forecast an improvement for year t in 86.9 percent of cases, compared to 61.3

TABLE 3
The Relation between Managers' Forecast Innovations and Forecast Errors in Year *t* and Those in Prior Years (NI Forecasts)

Panel A: Relation between Year *t* Forecast Innovations and Properties of Year $t-1$ Forecasts[a]

	Year $t-1$ Forecast Innovations		Row Total
Year $t-1$ Forecast Errors	**Innovations ≥ 0**	**Innovations < 0**	
Forecast Errors ≥ 0	8,299 (63.2% > 0)	2,807 (55.8% > 0)	11,106 (61.3% > 0)
Forecast Errors < 0	14,407 (87.5% > 0)	3,615 (84.6% > 0)	18,022 (86.9% > 0)
	22,706	6,422	29,128

Panel B: Relation between Year *t* Forecast Errors and Properties of Year $t-1$ Forecasts[a]

	Year $t-1$ Forecast Innovations		Row Total
Year $t-1$ Forecast Errors	**Innovations ≥ 0**	**Innovations < 0**	
Forecast Errors ≥ 0	7,243 (52.8% > 0)	2,343 (54.3% > 0)	9,586 (53.2% > 0)
Forecast Errors < 0	13,244 (31.4% > 0)	3,225 (32.9% > 0)	16,469 (31.7% > 0)
	20,487	5,568	26,055

(*continued on next page*)

TABLE 3 (continued)

Panel C: The Relation between Number of Past Forecast Errors and Properties of Current Period Forecasts[b]

		Year t Forecast Innovations			Year t Forecast Errors		
		Negative	**Positive**	**Total**	**Negative**	**Positive**	**Total**
Number of Positive Forecast Errors in Prior Three Years	None	617	5,313	5,930	3,788	1,683	5,471
	%	10.4%	89.6%		69.2%	30.8%	
	Once	1,786	5,281	7,067	3,587	2,700	6,287
	%	25.3%	74.7%		57.1%	42.9%	
	Twice	1,665	3,110	4,775	1,999	1,997	3,996
	%	34.9%	65.1%		50.0%	50.0%	
	Three times	800	1,209	2,009	527	902	1,429
	%	39.8%	60.2%		36.9%	63.1%	
	Total	4,868	14,913	19,781	9,901	7,282	17,183
	%	24.6%	75.4%		57.6%	42.4%	

---

[a] Cells show number of year t firm/years in each category, with number and percent of positive forecast innovations in year t shown in parentheses.

[b] For those firms with four consecutive years of forecast data, we tabulate the relation between the number of times in the prior three years that managers' initial forecasts result in positive forecast errors (0, 1, 2, or 3 times) and the sign of current period forecast innovations and forecast errors.

percent of cases when the year $t-1$ forecast error is non-negative (total of 11,106 observations). Thus, managers who issue forecasts that were too optimistic in year $t-1$ are substantially more likely to again issue overly optimistic forecasts in year t than managers who were not too optimistic.

Further, managers whose forecasts are too optimistic in year $t-1$ are more likely to issue year t forecasts that also result in negative forecast errors (68.3 percent of 16,469 observations) than managers whose forecasts are too pessimistic in year $t-1$ (46.8 percent of 9,586 observations). Thus, managers' tendency to be overly optimistic is persistent. These results are inconsistent with the reputation argument, which predicts negative rather than positive autocorrelation in forecast bias.

The data in Panels A and B of Table 3 relate managers' year $t-1$ forecasts to those in year t. To further assess the persistence of managers' forecast errors, Panel C of Table 3 provides evidence on the relation between three prior years of forecast errors and forecasts in year t. We perform these tests for firms with at least three consecutive years of prior data (results are similar when we consider two and four prior years of earnings data). We again report data for NI; results for Sales (not reported) are similar.

The data in Table 3, Panel C provide clear evidence of the persistence of managerial optimism. For managers of firms that consistently issue forecasts that turn out to be conservative (three prior years of positive forecast errors), year t forecast innovations are positive 60.2 percent of the time. As we move from this group to managers who are progressively more optimistic (with one, two, and three years of negative forecast errors in the three-year period), the fraction of positive innovations increases to 65.1 percent, 74.7 percent, and 89.6 percent, respectively. Thus, managers of firms who are consistently most over-optimistic in prior years are also the most optimistic in the current year. Even more strikingly, managers in these four groups had year t forecasts that proved too optimistic in 36.9 percent, 50.0 percent, 57.1 percent, and 69.2 percent of cases, respectively, which clearly show the tendency for more persistent past optimism to result in a greater likelihood of current period optimism. These results are hard to understand in a world where reputation effects penalize managers for forecast optimism.

We present two additional analyses of managers' initial forecasts. First, to better understand the economic forces that underlie managers' forecast optimism, we present cross-sectional regressions of the determinants of managerial forecast optimism. Second, we provide evidence on the relation between the accuracy of managers' past forecasts and the stock price response to the forecast news. If stock market participants understand that managers' forecast optimism is persistent, then we expect them to discount forecasts of managers with a track record of over-optimism.

We provide evidence on the cross-sectional determinants of forecast optimism using the logit regression specifications that we report in Table 4. Following the arguments in Section II, we include independent variables that capture:

(1) Firms' past earnings performance, which we expect is negatively related to forecast optimism (H2). As in Table 2, this variable is measured as ROA in period $t-1$. We also interact this variable with a year $t-1$ loss dummy, to investigate whether the relation between managers' forecast optimism and firm performance changes when their firms report losses.

(2) A dummy variable (*Prior Optim.*) set to 1 if the firm's initial forecast in year $t-1$ resulted in a negative forecast error (H6). We expect this variable to be positively related to optimism in year t given our results on the persistence of managerial optimism.

TABLE 4
Logit Regressions of the Determinants of Forecast Optimism and Associated Descriptive Statistics

Panel A: Descriptive Statistics for Regression Independent Variables (obs. = 30,580 firm/years)

Variable	Mean	Median	Standard Deviation	Minimum	Maximum
Lagged ROA	0.013	0.015	0.039	−0.221	0.103
Prior Optim.	0.537	1.000	0.499	0.000	1.000
Size (log total assets)	10.600	10.461	1.305	7.555	14.442
Insider Ownership (%)	8.080	1.946	11.464	0.013	51.353
Financial Inst. Own. (%)	21.958	19.518	13.186	0.213	55.707
Foreign Own. (%)	5.142	2.102	6.814	0	35.481

Panel B: Correlation Matrix for Regression Independent Variables (obs. = 30,580 firm/years)

Variable	*Lagged ROA*	*Prior Optim.*	*Size*	*Insider Own.*	*Financial Inst. Own.*
Prior Optim.	−0.351				
Size (log total assets)	−0.008	−0.016			
Insider Ownership (%)	0.161	−0.033	−0.380		
Financial Inst. Own. (%)	0.028	−0.013	0.607	−0.354	
Foreign Own. (%)	0.257	−0.099	0.413	−0.059	0.328

Panel C: Logistic Regressions of Forecast Optimism (measured dichotomously) on Hypothesized Determinants

	Optimism #1		*Optimism #2*	
	Coeff.	Z-stat.	Coeff.	Z-stat.
Intercept	2.45	13.8*	1.75	10.1*
Lagged ROA	−7.04	−8.94*	−14.70	−18.1*
*Loss * Lagged ROA*	6.41	5.70*	9.52	8.25*
Prior Optim.	0.52	16.2*	0.73	23.1*
Size	−0.10	−6.08*	−0.12	−7.75*
Insider Own.	0.004	3.15*	0.013	9.26*
Financial Inst. Own.	0.001	0.51	0.003	2.52*
Foreign Own.	−0.004	−1.42	0.011	4.16*
Year Effects	Yes		Yes	
Industry Effects	Yes		Yes	

---

* Denotes statistical significance at the 1 percent level (two-tailed).

Variable Definitions:

Optimism #1 = 1 if the forecast error for year *t* is negative, and 0 otherwise;
Optimism #2 = 1 if the forecast error for year *t* is negative and the forecast innovation is positive, and 0 otherwise;
Lagged ROA = net income for year *t*−1 divided by total assets at *t*−1 year-end;
Loss = 1 if net income in year *t*−1 is negative, and 0 otherwise;
Prior Optim. = dummy variable set to 1 if the forecast error in year *t*−1 is negative, and 0 otherwise;
Size = log of total assets at year-end *t*;
Insider Ownership = percentage ownership interest of management and board members;

(*continued on next page*)

TABLE 4 (continued)

Financial Inst. Own. = percentage ownership interest of financial institutions (banks and insurance companies); and

Foreign Own. = percentage ownership of foreign investors.

Outliers at or beyond the 1st and 99th percentiles of the distributions of the independent variables have been removed.

(3) Ownership structure. As discussed in Section II (H3), we expect firms whose shares are held by insiders and/or other main bank (*keiretsu*) entities to have managers who are less subject to external discipline and so more likely to issue optimistic forecasts. We use three ownership variables: (a) insider ownership (the fraction of shares held by management and/or directors), (b) ownership by financial institutions (which we expect mainly comprises holdings by financial institutions within each group), (c) ownership by foreign investors.

(4) Firm size. We include firm size as a control variable and because firm size may be negatively related to forecast optimism if larger firms are subject to greater external discipline (they may be cross-listed on overseas exchanges, or may be generally more visible) and/or are more likely to have managers who bear relatively larger reputational costs.

We use two dichotomous measures of forecast optimism as the dependent variable in these regressions: (1) a variable set to 1 if the forecast error in year t is negative, and 0 otherwise; (2) a variable set to 1 if the forecast innovation for year t is positive and the forecast error for year t is negative, and 0 otherwise. Because of the systematic year-to-year variation in forecast optimism (Figure 1), we include year fixed effects as well as industry fixed effects.

We report descriptive statistics for the regression variables, along with their correlations, in Table 4 (Panels A and B). These statistics show that ownership by financial institutions is large (mean 22.0 percent, median 19.5 percent) compared to insider ownership (mean 8.1 percent, median 1.9 percent) and foreign ownership (mean 5.1 percent, median 2.1 percent). In terms of correlations (Panel B), lagged ROA is negatively related to *Prior Optim.* (correlation of −0.351), consistent with our earlier finding that optimism is negatively related to contemporaneous performance. Lagged ROA is positively related to insider ownership (correlation of 0.161) and foreign ownership (0.257) but essentially unrelated to ownership by financial institutions. *Size* is negatively related to insider ownership (−0.380), consistent with evidence from other countries, and positively related to ownership by financial institutions (0.607) and foreign investors (0.413). Ownership by financial institutions is positively related to foreign ownership (0.328) and negatively related to insider ownership (−0.354).

The regression results in Table 4, Panel C indicate that these variables are generally significantly related to forecast optimism in the ways we expect. *Lagged ROA* is significantly negatively related to both measures of optimism (Z-statistics of −8.94 and −18.1), consistent with the results in Table 2. Because these regressions control for year effects and the persistence of past optimism, this shows that the Table 2 results are not attributable to either particular sample years or the persistence result. We also find that the *Lagged ROA/Loss* interaction variable is significantly positive, indicating that managers are relatively more optimistic when their firms report losses, perhaps because they expect the loss will be reversed or are otherwise unwilling to forecast another loss. The *Prior Optim.* variable is positive and highly significant in these regressions (Z-statistics of 16.2 and 23.1), confirming our earlier findings that managers of these firms are persistently optimistic. This result holds after controlling for both year effects and for the negative cross-sectional relation between optimism and firm performance.

The results for the ownership variables are broadly consistent with our argument that managers of firms with larger insider and/or group cross-holdings are more likely to be optimistic because they face less external discipline. The coefficient on the insider ownership variable (which is most clearly associated with management entrenchment) is positive and significant at the 1 percent level in both regressions while that on the financial institutions ownership variable is significantly positive in one of the two regressions. Results for the foreign ownership variable are mixed (significantly positive in one regression and insignificant in the other). Firm size is strongly negatively related to forecast optimism, indicating that managers of larger firms tend to issue more conservative forecasts, other factors held constant. This suggests that managers of larger firms may worry more about reputation effects.

Overall, the regression results in Table 4 confirm the robustness of our earlier findings that forecast optimism is both persistent and negatively related to firm performance. In addition, there is evidence that managers of smaller firms and firms with larger insider ownership are more optimistic.

We report on the information content tests in Table 5. The dependent variable in these regressions is the market-adjusted stock return cumulated over the five trading days around the forecast release date (days -3 to $+1$), using the Japanese Topix index to measure market returns. As discussed in Section II, the initial earnings forecast for year t is released concurrently with actual earnings for year $t-1$ as well as the new dividend amount. Consequently, to assess the information content of the forecast news, we regress the announcement period return on an earnings surprise measure (realized earnings minus the most recent management earnings forecast), a dividend surprise measure (announced dividend minus previous dividend), and the forecast innovation.[18] We cannot compute the earnings surprise variable (which relies on management forecast revisions) until fiscal 2002, the first year for which we have forecast revisions; dividend surprises are also unavailable until this year.

Table 5 first reports simple regressions of the announcement period returns on the forecast innovations for the full sample period (1997–2007) as well as for the 2002–2007 subperiod for which we can compute earnings surprises (rows (1) and (2)). These simple regressions reveal that the initial forecasts are informative: the coefficient on the forecast innovation is positive and statistically significant for both the full period and the 2002–2007 subperiod (coefficients of 0.124 and 0.156, with t-statistics of 16.3 and 14.9, respectively). The fact that the coefficients are similar across subperiods provides some assurance that the 2002–2007 subperiod for which we can estimate the full specification is not unusual.

The next regression specification (row (3)) shows that the forecast innovations are still informative after we control for the earnings and dividend surprises (2002–2007 period only). Coefficients on both surprise variables are positive and statistically significant, as expected, while the coefficient on the forecast innovation is somewhat larger than in the simple regression (coefficient of 0.204, t = 10.8). The fact that the forecast innovation

---

[18] The earnings news and forecast innovation variables are deflated by market value of equity (per share earnings and forecast data are not available in Japan) while the dividend variable (which is available on a per share basis) is deflated by stock price.

TABLE 5
Regressions of Announcement Period Returns on Forecast, Earnings, and Dividend Surprise Variables: Managers' Initial Earnings Forecasts from 1997–2007

	Sample Period	Obs.	Intercept	*Forecast Innov.*	*Earnings Surprise*	*Dividend Surprise*	*Innov.* Prior Optim.*	*Innov.* Q2*	*Innov.* Q3*	*Innov.* Q4*	*Innov.* Q5*	Adjusted R^2
(1)	1997–2007	26,572	0.007 (13.3*)	0.124 (16.3*)								0.021
(2)	2002–2007	16,780	0.002 (3.02*)	0.156 (14.9*)								0.029
(3)	2002–2007	9,971	0.000 (.56)	0.204 (10.8*)	0.955 (7.03*)	2.263 (7.18*)						0.047
(4)	2002–2007	9,971	−0.002 (−1.89)	0.096 (6.03*)	1.098 (8.83*)	1.896 (6.27*)		0.567 (7.61*)	1.179 (8.13*)	0.985 (9.27*)	0.937 (9.20*)	0.129
(5)	2002–2007	9,971	0.001 (1.64)	0.817 (10.3*)	0.981 (7.33*)	2.053 (6.67*)	−0.675 (−8.32*)					0.080

Regression dependent variable is five-day market-adjusted stock return centered on forecast release date (t-statistics in parentheses). We use returns on the Japanese Topix index to measure market returns.

Variable Definitions:

Initial Forecast Innovation = management forecast for year t minus actual earnings for year $t-1$;
Earnings Surprises = realized earnings for year $t-1$ minus most recent management forecast of year $t-1$ earnings;
Dividend Surprise = announced change in dividends per share;
Prior Optim. = dummy variable set to 1 if managers' initial earnings forecast in year $t-1$ resulted in negative forecast error, and 0 otherwise; and
Qk = dummy variable set to 1 if firm's *ROA* in year $t-1$ was in *k*th quintile, and 0 otherwise, Quintile 1 firms with lowest *ROA* in year $t-1$.

t-statistics are White (1980) heteroscedasticity consistent t-statistics.

variable is significant in these regressions is evidence that managers' initial annual forecasts are informative to market participants in spite of the general optimistic bias.

We next examine whether the information content of the forecast innovations declines as likely forecast bias increases and the forecast becomes less credible (H5). We test this in two ways. First, we examine whether the coefficient on the forecast innovation varies across the profitability quintiles used in Table 2. Because the extent of the optimistic bias is inversely related to profitability, we expect the market's response to forecast news is also inversely related to profitability. Second, we examine whether the stock market response to forecasts depends on whether the manager's previous forecast was overly optimistic (there was a negative forecast error). To do this, we include an interaction term in the regression, obtained by multiplying the forecast innovation by a dummy variable (*Prior Optim.*) set to 1 if the manager's initial forecast in the prior year results in a negative forecast error, and 0 otherwise.

The results support the prediction that the market response is smaller for less credible forecasts. First, when we examine how the coefficient on forecast innovations varies across the ROA quintiles (row (4)), the coefficient for Quintile 1 is lowest, at 0.096 (t = 6.03). The incremental coefficient for Quintile 2 is 0.567 (t = 7.61), indicating that it is significantly larger than that for Quintile 1. The coefficients for Quintile 3 through 5 are also positive, at 1.179, 0.985, and 0.937, respectively, and statistically significant (t-statistics are 8.13, 9.27, and 9.20, respectively) indicating that these coefficients are also larger than that for Quintile 1. These results indicate that the market response is smaller for Quintile 1, where optimism is strongest, than for any of the other quintiles, which supports our prediction that investors' reaction to forecasts depends on their likely credibility.

The second credibility specification (row (5)) also supports the hypothesis that the market's response to forecast news is lower for less credible forecasts. In this regression the coefficient on the forecast innovation variable is 0.817 (t = 10.3) while that on the interaction with prior optimism is -0.675 (t = -8.32). Thus, when managers' prior forecasts result in a negative forecast error (were overly optimistic), investors discount managers' year *t* forecasts to a substantial extent.

Overall, the regressions in Table 5 provide evidence that initial management forecasts in Japan are, on average, informative (credible) to market participants but that they discount those forecasts that are most likely to be biased.

Evidence on Forecast Revisions

To this point, we have analyzed evidence on managers' initial forecasts for each fiscal year. Our next set of tests uses data on management forecast revisions, which are available for 2002–2007. We first summarize these data in Panel A of Table 6. The columns represent the number of forecast observations in each year that are initial forecasts, first revisions, second revisions, etc. (revisions includes forecasts that confirm the initial forecast). Apart from 2007, for which we do not have a full year of data, the numbers in the "1st" column (which denotes the first revision) are close to those in the initial forecast column—99.6 percent of initial forecasts are accompanied by at least one revision. This is consistent with the Timely Disclosure Rules. The numbers in the subsequent columns indicate that most managers issue additional revisions; managers provide a second revision in 87.8 percent of cases, a third revision in 65.5 percent of cases, a fourth revision in 30.6 percent of cases,

TABLE 6
Timing and Relative Frequency of Management Earnings Forecast Revisions

Panel A: Frequency of Initial Forecasts and Subsequent Forecast Revisions[a]

			Revisions										
Year	**Total**	**Initial**	**1st**	**2nd**	**3rd**	**4th**	**5th**	**6th**	**7th**	**8th**	**9th**	**10th**	**11th**
2002	9,440	3,114	3,105	2,096	929	162	27	6	1	0	0	0	0
2003	12,716	3,195	3,181	2,792	2,161	1,020	316	48	3	0	0	0	0
2004	14,297	3,283	3,274	3,081	2,729	1,399	468	52	9	2	0	0	0
2005	14,921	3,360	3,351	3,186	2,865	1,531	523	81	16	3	3	1	1
2006	15,358	3,455	3,445	3,257	2,959	1,592	551	74	19	5	1	0	0
2007	12,554	3,653	3,618	3,201	1,488	438	131	19	6	0	0	0	0
Total	79,286	20,060	19,974	17,613	13,131	6,142	2,016	280	54	10	4	1	1
		100.0%	99.6%	87.8%	65.5%	30.6%	10.0%	1.4%	0.3%	0%	0%	0%	0%

Panel B: Intra-Year Timing of Management Forecasts[b]

			Revisions										
Forecast Release Type	**Total**	**Initial**	**1st**	**2nd**	**3rd**	**4th**	**5th**	**6th**	**7th**	**8th**	**9th**	**10th**	**11th**
Stand-Alone	19,089	69	4,672	5,631	3,284	3,347	1,787	241	46	7	4	1	
	24.1%	0.3%	23.4%	32.0%	25.0%	54.5%	88.6%	86.1%	85.2%	70.0%	100.0%	100.0%	
Annual	19,956	19,933	13	2	1	5	2						
	25.2%	99.4%	0.1%	0.0%	0.0%	0.1%	0.1%						
Semi-Annual	19,904	21	4,781	10,987	3,840	248	20	5	2				
	25.1%	0.1%	23.9%	62.4%	29.2%	4.0%	1.0%	1.8%	3.7%	0.0%			
1st Quarter	10,801	16	10,492	279	13		1						
	13.6%	0.1%	52.5%	1.6%	0.1%		0.0%						
3rd Quarter	9,536	21	16	714	5,993	2,542	206	34	6	3			1
	12.0%	0.1%	0.1%	4.1%	45.6%	41.4%	10.2%	12.1%	11.1%	30.0%			100.0%
Total	79,286	20,060	19,974	17,613	13,131	6,142	2,016	280	54	10	4	1	1

---

[a] Panel A summarizes the frequency of initial management forecasts and subsequent revisions. For fiscal 2007, we only have data through December 2007.

[b] Panel B summarizes the frequency of management forecasts, including revisions, categorized by the type of release, which is either "stand alone" (forecast released separate from earnings announcement) or as the earnings announcement at which the forecast is released. For fiscal 2007, we only have data through December 2007.

and a fifth revision forecast in 10.0 percent of cases, with managers of a few firms revising more often than this.[19]

Panel B of Table 6 provides information on when management provides earnings forecasts. As expected, the large majority of initial forecasts are released at the annual earnings announcement date. Most of the subsequent revisions are also released on earnings announcement dates; for the first revisions, for example, 52.5 percent are released with first quarter earnings and 23.9 percent are released with semi-annual earnings (23.4 percent are released on a stand-alone basis). Similarly, 62.4 percent of the second revisions are released with semi-annual earnings and 32.0 percent are released on a stand-alone basis, while 45.6 percent of third revisions are released with third quarter earnings, 29.2 percent with semi-annual earnings, and 25.0 percent on a stand-alone basis.

We next provide evidence on how managers revise their forecasts. To conduct these tests, we divide the fiscal year into six periods labeled as follows:

T1: forecasts released at the annual earnings announcement (initial forecasts);
T2: forecasts released after T1 through the first quarter's earnings announcement date;
T3: forecasts released after T2 through the second quarter's earnings announcement date;
T4: forecasts released after T3 through the third quarter's earnings announcement date;
T5: forecasts released after T4 through the end of the fiscal year; and
T6: forecasts released after the end of the fiscal year and before the annual earnings announcement date (pre-announcements).

Table 7 reports on the forecast revisions and forecast errors for forecasts issued during each of these periods. In computing these forecast errors, we do not update for earnings realizations that occur during the year, even though managers' forecasts are likely to reflect year-to-date realizations and so will naturally decline as the fiscal year unfolds. We do this because our objective is to indicate descriptively how managers' forecasts evolve during the year, rather than computing a forecast error for each sub-period within the year. Thus, a declining forecast error can arise because managers' optimism decreases, because managers receive information about year-to-date realized earnings, or both.

The results in Table 7 show that managers generally revise their forecasts downward during the fiscal year, as expected. First, consistent with Table 2, we see that for initial forecasts (T1), forecast innovations are systematically positive while forecast errors are systematically negative. These results are not as strong as those in Table 2 because they exclude the earlier years (1997–2001) when managers' forecasts were most optimistic (Figure 1). The large majority of forecasts issued during T2 (which includes the first quarter earnings announcement) are confirmatory forecasts; for the NI forecasts, for example, 68.4 percent of the T2 revisions are zero. Similar to the initial forecasts, these forecasts are optimistically biased, with a majority of negative forecast errors (54.1 percent negative versus 45.6 percent positive for the NI forecasts).

Second, after T2, managers' revisions are no longer optimistically biased and, in fact, tend to be conservative. For both Sales and NI, the ratio of positive to negative forecast errors increases as we move from T3 through T6 (for NI the ratios are: 53.2 percent to 46.5 percent, 58.0 percent to 41.4 percent, 69.1 percent to 28.5 percent, and 49.5 percent

---

[19] The number of forecast revisions in Japan is more numerous than in the U.S. Anilowski et al. (2007) find that for the set of firm/years when annual earnings forecasts are released, managers release a single forecast about half the time (51 percent of observations), two forecasts 18 percent of the time, three forecasts 11 percent of the time, four forecasts 8 percent of the time, and five or more forecasts in the remaining 12 percent of cases. Ajinkya et al. (2005) also provide evidence on the frequency of multiple forecasts in U.S. data.

TABLE 7
Properties of Management Forecasts, Including Revisions, Partitioned According to When Released during Fiscal Year

Panel A: Sales

	T1		T2		T3		T4		T5		T6	
Forecast Innovations/Revisions												
Mean	0.075***		−0.002***		−0.000		−0.006***		−0.015***		−0.005***	
Median	0.038***		0.000***		0.000***		0.000***		0.000***		0.000***	
Obs.	19,987		21,692		17,831		12,573		1,041		6,162	
Positive	16,113	80.6%	2,636	12.2%	4,353	24.4%	1,418	11.3%	277	26.6%	2,773	45.0%
Confirmatory	25	0.1%	15,499	71.5%	9,593	53.8%	9,145	72.7%	370	35.5%	422	6.8%
Negative	3,849	19.3%	3,557	16.4%	3,885	21.8%	2,010	16.0%	394	37.8%	2,967	48.1%
Forecast Errors												
Mean	−0.007***		−0.005***		−0.001		0.004***		0.009***		0.001*	
Median	−0.004***		−0.002***		0.001***		0.003***		0.003***		0.000***	
Obs.	16,792		19,072		13,698		12,098		1,008		6,163	
Positive	7,887	47.0%	9,076	47.6%	7,143	52.1%	6,838	56.5%	688	68.3%	3,394	55.1%
Zero	8	0.0%	15	0.1%	11	0.1%	21	0.2%	26	2.6%	1,896	30.8%
Negative	8,897	53.0%	9,981	52.3%	6,544	47.8%	5,239	43.3%	294	29.2%	873	14.2%

(continued on next page)

TABLE 7 (continued)

Panel B: Net Income

	T1		T2		T3		T4		T5		T6	
Forecast Innovations/Revisions												
Mean	0.019***		−0.002***		−0.003***		−0.004***		−0.014***		−0.010***	
Median	0.005***		0.000***		0.000***		0.000***		−0.001***		−0.001***	
Obs.	19,975		21,373		17,822		12,286		1,036		6,159	
Positive	15,372	77.0%	3,077	14.4%	4,321	24.2%	1,574	12.8%	320	30.9%	2,669	43.3%
Confirmatory	56	0.3%	14,609	68.4%	9,320	52.3%	8,324	67.8%	186	18.0%	226	3.7%
Negative	4,547	22.8%	3,687	17.3%	4,181	23.5%	2,388	19.4%	530	51.2%	3,264	53.0%
Forecast Errors												
Mean	−0.013***		−0.011***		−0.009***		−0.004***		−0.002		0.000	
Median	−0.001***		−0.001***		0.000***		0.001***		0.001***		0.000***	
Obs.	16,781		18,756		13,691		11,813		1,004		6,160	
Positive	7,624	45.4%	8,554	45.6%	7,278	53.2%	6,856	58.0%	694	69.1%	3,047	49.5%
Zero	37	0.2%	61	0.3%	51	0.4%	61	0.5%	24	2.4%	1,714	27.8%
Negative	9,120	54.3%	10,141	54.1%	6,362	46.5%	4,896	41.4%	286	28.5%	1,399	22.7%

---

*, **, *** Indicate statistical significance at 5 percent, 2.5 percent, and 1 percent, respectively, under two-tailed tests.

We divide the forecasts into six time periods according to when they are released during the fiscal year: T1 denotes forecasts released at the annual earnings announcement, T2 denotes forecasts released after T1 up through the first quarter earnings announcement date, T3 denotes forecasts released after T2 up through the second quarter earnings announcement date, T4 denotes forecast revisions released after T3 up through the third quarter earnings announcement date, T5 denotes forecast revisions released after T4 through the remainder of the fiscal year, T6 denotes forecast revisions released after the end of the fiscal year but before the annual earnings announcement date. For Fiscal 2007, we only have data through December 2007. For the initial management forecast in each fiscal year we compute forecast innovations as the forecast innovation (management's forecast for year *t* minus realized earnings for year $t-1$) divided by lagged total assets. For subsequent forecasts, we compute the revision as the newly released management forecast for year *t* minus management's previous forecast for the same period, all divided by lagged total assets. Forecast errors are defined as realized earnings for year *t* minus management forecast for the same period, divided by lagged total assets.

to 22.7 percent, respectively). Thus, managers tend to revise their forecasts to reverse the initial optimism. This is also evident in the revisions themselves; as the fiscal year goes by we see an increasing ratio of negative to positive revisions (e.g., for NI, 19.4 percent to 12.8 percent for T4, 51.2 percent to 30.9 percent for T5). For revisions made late in the fiscal year (T5) and for pre-announcements (T6), substantially more revisions are negative than is true earlier in the fiscal period, which is also evidence of managers correcting overly optimistic forecasts before earnings are announced. Notice the preponderance of positive forecast errors associated with these forecasts; at T5, 69.1 percent of forecasts result in positive forecast errors (versus 28.5 percent negative); at T6, 49.5 percent of forecasts result in positive forecast errors (versus 22.7 percent negative). These results suggest that managers' forecast revisions offset their initial forecast optimism and are typically conservative or unbiased.

Table 8 reports on earnings surprises, calculated as realized earnings for each firm/year minus the most recently available management forecast. Because we are interested in the extent to which managers' initial optimism is offset by subsequent revisions, we also present results for quintiles based on the magnitude of the initial forecast innovation, sorted from lowest (Quintile 1) to highest (Quintile 5).

The evidence in Table 8 indicates that, overall, managers revise their forecasts to avoid negative earnings surprises. For Sales, 68.3 percent of the surprises are non-negative (56.9 percent positive and 11.4 percent zero), while for NI 67.3 percent of the surprises are non-negative (56.5 percent positive and 10.8 percent zero). This asymmetry suggests that Japanese managers bear larger costs when earnings surprises are negative than when they are positive, consistent with results for the U.S. We also find a modest relation between managers' initial optimism (as measured by forecast innovations) and earnings surprises, with a greater tendency for managers who were more optimistic at the beginning of the fiscal period to report earnings surprises that result in negative earnings surprises. Overall, though, the impression is that managers act to reverse the systematic optimism evident for initial forecasts, so that most earnings surprises are either zero or positive.

Our final tests provide evidence on the information content of forecast revisions. To do this, we estimate regressions similar to those reported in Table 5, except that the innovation variable (the news associated with the initial forecast) is replaced with the forecast revision variable, which has the same interpretation (forecast news). For the revisions we do not include measures of concurrent earnings and dividend surprises because these data are not available. We report these results in Table 9.

The first regression in Table 9 is a simple regression of the announcement period return on the revision news. We find that the revisions are informative, on average: the coefficient on the revision is 0.508 and highly statistically significant ($t = 32.3$). This coefficient is substantially larger and more significant than that for the initial forecast innovations in Table 5 (0.156, $t = 14.9$). Similar to the Table 5 regressions, we also report a specification that allows the revision coefficient to vary across the lagged ROA quintiles. As was the case in Table 5, the coefficient for Quintile 1 is significantly lower than that for the other quintiles. This is evidence that market participants find these revisions to be less credible than those in the other quintiles. However, note that the coefficient magnitudes are quite different to those in Table 5. The coefficient for Quintile 1 is 0.360 (versus 0.096 in Table 5) and the incremental coefficient for Quintile 2 is 0.078 with a t-statistic of 2.1 (versus 0.567, $t = 7.6$). Thus, in relative terms the discount applied to the Quintile 1 forecasts is smaller than in Table 5, consistent with the prediction that the revisions are more credible than the original forecasts.

TABLE 8
Earnings Surprises Overall and by Quintile Based on Initial Forecast Innovation

Panel A: Sales

	Overall		Quintile 1 (lowest innov.)		Quintile 2		Quintile 3		Quintile 4		Quintile 5 (highest innov.)	
Mean	0.003***		0.005***		0.004***		0.005***		0.001**		0.001	
Median	0.000***		0.000***		0.000***		0.000***		0.000***		0.000***	
Obs.	16,858		3,373		3,372		3,373		3,372		3,368	
Positive	9,586	56.9%	2,044	60.6%	1,986	58.9%	1,969	58.4%	1,830	54.3%	1,757	52.2%
Zero	1,918	11.4%	455	13.5%	354	10.5%	332	9.8%	358	10.6%	419	12.4%
Negative	5,354	31.7%	874	25.9%	1,032	30.6%	1,072	31.8%	1,184	35.1%	1,192	35.4%

Panel B: Net Income

	Overall		Quintile 1 (lowest innov.)		Quintile 2		Quintile 3		Quintile 4		Quintile 5 (highest innov.)	
Mean	−0.000*		0.000*		0.000		0.000*		0.000*		−0.003***	
Median	0.000***		0.000***		0.000***		0.000***		0.000***		0.000***	
Obs.	16,750		3,360		3,343		3,349		3,350		3,348	
Positive	9,463	56.5%	1,983	59.0%	2,015	60.3%	1,990	59.4%	1,848	55.2%	1,627	48.6%
Zero	1,802	10.8%	361	10.7%	329	9.8%	292	8.7%	350	10.4%	470	14.0%
Negative	5,485	32.7%	1,016	30.2%	999	29.9%	1,067	31.9%	1,152	34.4%	1,251	37.4%

---

*, **, *** Show statistical significance at 5 percent, 2.5 percent, and 1 percent, respectively.

The table reports the distributions of earnings surprises from the latest management forecast revisions. Surprise is defined as [realized earnings − latest management forecast]/total assets at the end of prior year-end. Overall sample is sorted into quintiles based on the magnitude of innovation at initial management forecast. Partition formation is made year by year and then pooled into each quintile. For Fiscal 2007, we only have data through December 2007.

TABLE 9

Regressions of Announcement Period Stock Returns on Forecast Innovations: Managers' Forecast Revisions from 1997–2007

Sample Period	Obs.	Intercept	*Forecast Innov.*	*Innov.* Prior Optim.*	*Innov.* Q2*	*Innov.* Q3*	*Innov.* Q4*	*Innov.* Q5*	Adjusted R^2
2002–2007	50,911	0.002 (5.3*)	0.508 (32.3*)						0.047
2002–2007	50,911	0.002 (5.3*)	0.360 (17.8*)		0.078 (2.1*)	0.291 (6.0*)	0.593 (7.7*)	0.722 (9.0*)	0.058
2002–2007	50,911	0.000 (.57)	0.994 (21.7*)	−0.617 (−12.6*)					0.058

---

* Indicates statistical significance at 5 percent level.

Regression dependent variable is five-day market-adjusted stock return centered on forecast release date (t-statistics in parentheses). We use returns on the Japanese Topix index to measure market returns. For Fiscal 2007, we only have data through December 2007.

Variable Definitions:

Forecast Innov. = managers' forecast revisions (management's revised forecast for year t minus management's previous forecast for year t) deflated by lagged total; and

Prior Optim. = dummy variable set to 1 if managers' initial earnings forecast in year $t-1$ resulted in negative forecast error, and 0 otherwise; and

Qk = dummy variable set to one if firm's *ROA* in year $t-1$ was in *k*th Quintile, and 0 otherwise, where Q1 contains firm/years with the lowest *ROA*.

t-statistics are White (1980) heteroscedasticity-consistent t-statistics.

Similarly, when we estimate the specification with the *Prior Optim.* interaction, it is again significant, suggesting that the revisions are less credible when management has been overly optimistic in the prior period. In addition, the magnitude of the unconditional coefficient on forecast news here is 0.377 (0.994 – 0.617) which is larger than that the corresponding coefficient in Table 5 of 0.142 (0.817 – 0.675). This again suggests that revisions are more credible than the initial forecasts, although we are do not test the statistical significance of this difference. Overall, the evidence supports our prediction that managers' revisions are more informative than their initial forecasts.

V. SUMMARY AND CONCLUSIONS

This study provides evidence on the properties of management forecasts in Japan, where securities laws and stock exchange listing requirements effectively require that managers issue earnings forecasts but where managers have considerable latitude over the numbers they release. This makes management forecasting in Japan markedly different from that in other countries, such as the U.S., where management forecasts are voluntary and managers face significant costs if they bias their forecasts.

Using a comprehensive sample of Japanese firms, we find that managers' initial annual earnings forecasts are, on average, upward biased (optimistic). This phenomenon is more pronounced for firms whose past earnings performance is poor, firms with higher levels of insider/managerial ownership, and smaller firms. We also find that managers' forecast optimism is highly persistent from one year to the next, which is surprising if market forces discipline forecast accuracy.

When we examine the information content of managers' initial forecasts for the fiscal year, we find that forecast news is associated with announcement period stock returns, on average, indicating that these forecasts are informative. However, we also find that forecasts are less informative for firms with poor performance and for managers whose prior forecasts were overly optimistic, suggesting that market participants discount forecasts that are likely to be less credible.

During the fiscal year, Japanese managers revise their annual earnings forecasts to undo their initial optimism, so that by the end of the year investors' expectations are, on average, in line with subsequent realizations. There is some evidence that these revisions are conservative in that most (around two-thirds) of earnings surprises are non-negative, suggesting that negative earnings surprises may be more costly to Japanese managers. Consistent with these forecasts revisions being less biased, we find that the stock price response to forecast news contained in the revisions is larger than that to the initial forecasts.

The credibility of managers' earnings forecast has been a central issue in the disclosure literature for several decades. Yet most of the evidence we have on this phenomenon is drawn from settings where forecasts are voluntary, which makes it hard to generalize the results because there is likely to be information in managers' *decision* to forecast as well as in the forecast news itself. The Japanese setting allows us to investigate the properties of management earnings forecasts in a regime where forecast disclosures are effectively mandated, and so provides evidence on how the properties of management forecasts change when they are no longer voluntary.

REFERENCES

Ajinkya, B. B., and M. J. Gift. 1984. Corporate managers' earnings forecasts and symmetrical adjustments of market expectations. *Journal of Accounting Research* 22: 425–444.

———, S. Bhoraj, and P. Sengupta. 2005. The association between outside directors, institutional investors and the properties of management earnings forecasts. *Journal of Accounting Research* 43: 343–376.

Anilowski, C., M. Feng, and D. J. Skinner. 2007. Does earnings guidance affect market returns? The nature and information content of aggregate earnings guidance. *Journal of Accounting and Economics* 44: 36–63.

Aoki, M., H. Patrick, and P. Sheard. 1994. The Japanese main bank system: An introductory overview. In *The Japanese Main Bank System*, edited by Aoki, M., and H. Patrick. London, U.K.: Oxford University Press.

Atiase, R. K., H. Li, S. Supattarakul, and S. Y. Tse. 2005. Market reaction to multiple contemporaneous signals: Earnings announcements and future earnings guidance. *Review of Accounting Studies* 10: 497–525.

Ben-David, I., J. R. Graham, and C. Harvey. 2007. Managerial overconfidence and corporate policies. Working paper, Duke University and University of Chicago.

Brooks, L. D., and D. A. Buckmaster. 1976. Further evidence of the time series properties of accounting income. *The Journal of Finance* 31: 1359–1373.

Conroy, R. M., K. Eades, and Y. S. Park. 1998. Fundamental information and stock prices in Japan: Evidence from earnings surprises and management predictions. *International Journal of Forecasting* 14: 227–244.

Conroy, R., K. Eades, and R. S. Harris. 2000. A test of the relative pricing effects of dividends and earnings: Evidence from simultaneous announcements in Japan. *The Journal of Finance* 40: 1199–1227.

Cooper, R. 1995. *When Lean Enterprises Collide: Competing through Confrontation.* Boston, MA: Harvard Business School Press.

———, and R. S. Kaplan. 1998. *Cost and Effect: Using Integrated Cost Systems to Drive Profitability and Performance.* Boston, MA: Harvard Business School Press.

Darrough, M. N., and T. S. Harris. 1991. Do management forecasts of earnings affect stock prices in Japan? In *Japanese Financial Market Research*, edited by W. T. Ziemba, W. Bailey, and Y. Hamao. Amsterdam, The Netherlands: North Holland.

Fama, E. F., and M. C. Jensen. 1983. Separation of ownership and control. *The Journal of Law & Economics* 26: 301–325.

———, and K. R. French. 2000. Forecasting profitability and earnings. *Journal of Business* 73: 161–175.

Freeman, R. N., J. A. Ohlson, and S. H. Penman. 1982. Book rate-of-return and prediction of earnings changes: An empirical investigation. *Journal of Accounting Research* 20: 639–653.

Ginsburg, T., and G. Hoetker. 2006. The unreluctant litigant? An empirical analysis of Japan's turn to litigation. *The Journal of Legal Studies* 35: 31–59.

Hoshi, T., and A. K. Kashyap. 2001. *Corporate Financing and Governance in Japan: The Road to the Future.* Cambridge, MA: The MIT Press.

Hutton, A. P., and P. C. Stocken. 2006. Effect of reputation and the credibility of management forecasts. Working paper. Dartmouth College.

Kahneman, D., and D. Lovallo. 1993. Timid choices and bold forecasts: A cognitive perspective on risk taking. *Management Science* 39: 17–31.

Kaplan, S. N. 1994. Top executive reward and firm performance: A comparison of Japan and the United States. *The Journal of Political Economy* 102: 510–546.

Kasznik, R., and B. Lev. 1995. To warn or not to warn: Management disclosures in the face of an earnings surprise. *The Accounting Review* 70: 113–134.

King, R., G. Pownall, and G. Waymire. 1990. Expectations adjustment via timely expectations forecast disclosure: Review, synthesis, and suggestions for future research. *Journal of Accounting Literature* 9: 113–144.

Lakonishok, J., A. Shleifer, and R. W. Vishny. 1994. Contrarian investment, extrapolation, and risk. *The Journal of Finance* 49: 1541–1578.

Leone, A., and S. Rock. 2002. Empirical tests of budget ratcheting and its effects on managers' discretionary accrual choices. *Journal of Accounting and Economics* 33: 43–67.

Lev, B., and S. H. Penman. 1990. Voluntary forecast disclosure, nondisclosure, and stock prices. *Journal of Accounting Research* 28: 49–76.

Malmendier, U., and G. Tate. 2005. CEO overconfidence and corporate investment. *The Journal of Finance* 60: 2661–2700.

Markham, J. 2003. Super-regulator: A comparative analysis of securities and derivatives regulation in the United States, Great Britain, and Japan. *Brooklyn Journal of International Law* 28: 319–410.

McNichols, M. 1989. Evidence of informational asymmetries from management earnings forecasts and stock returns. *The Accounting Review* 64: 1–27.

Morck, A., A. Shleifer, and R. W. Vishny. 1988. Management ownership and market valuation: An empirical analysis. *Journal of Financial Economics* 20: 293–315.

Patell, J. M. 1976. Corporate forecasts of earnings per share and stock price behavior: Empirical tests. *Journal of Accounting Research* 14: 246–276.

Penman, S. H. 1980. An empirical investigation of the voluntary disclosure of corporate earnings forecasts. *Journal of Accounting Research* 18: 132–160.

Pownall, G., and G. Waymire. 1989. Voluntary disclosure credibility and securities prices: Evidence from management earnings forecasts. *Journal of Accounting Research* 27: 227–245.

Rogers, J. L., and P. Stocken. 2005. Credibility of management forecasts. *The Accounting Review* 80: 1233–1260.

Shleifer, A., and R. W. Vishny. 1986. Large shareholders and corporate control. *The Journal of Political Economy* 94: 461–488.

Skinner, D. J. 1994. Why firms voluntarily disclose bad news. *Journal of Accounting Research* 32: 38–60.

Stocken, P. C. 2000. Credibility of voluntary disclosure. *The Rand Journal of Economics* 31: 359–374.

Verrecchia, R. E. 1983. Discretionary disclosure. *Journal of Accounting and Economics* 5: 179–194.

West, M. D. 2001. Why shareholders sue: The evidence from Japan. *The Journal of Legal Studies* 30: 351–382.

Zimmerman, J. L. 2006. *Accounting for Decision Making and Control.* 5th edition. New York, NY: McGraw-Hill Irwin.

THE ACCOUNTING REVIEW
Vol. 84, No. 5
2009
pp. 1607–1638

American Accounting Association
DOI: 10.2308/accr.2009.84.5.1607

A Modified Audit Production Framework: Evaluating the Relative Efficiency of Audit Engagements

W. Robert Knechel
University of Florida, Maastricht University, and The University of Auckland

Paul Rouse
The University of Auckland

Caren Schelleman
Maastricht University

ABSTRACT: We develop a model of audit production based on Data Envelopment Analysis (DEA) using labor cost as input and hours spent on evidence-gathering activities that determine the level of assurance as output. Client characteristics are considered exogenous factors that affect audit production as a whole. We apply the model to a sample of U.S.-based engagements from an international accounting firm. Results indicate that a constrained DEA model using variable returns to scale is appropriate for modeling audit production. We find that audits are more efficient for clients that are larger, have a December year-end, and are highly automated. Audits are less efficient when the auditor relies on internal control, tax services are provided, and the client has subsidiaries. We also find that a well-specified regression-based production model can control for factors that influence auditor efficiency. Finally, we find that inefficiencies are impounded in fees for some industries and firm offices.

Keywords: *audit production; Data Envelopment Analysis; efficiency.*

Data Availability: *The data used in this study are proprietary to the audit firm studied and cannot be released by the authors.*

This paper benefited from comments by participants at the 2004 European Accounting Association Annual Congress in Prague, the 2005 Auditing Section Midyear Conference in New Orleans, and workshop participants at Maastricht University. We especially thank Mike Stein for his extremely helpful comments at the 2005 Auditing Section Midyear Conference. We also acknowledge the late Mathew Tsui for his involvement in the early stages of this study.

Editor's note: Accepted by Steven Kachelmeier, with thanks to Bill Kinney for serving as editor on a previous version.

Submitted: January 2006
Accepted: December 2008
Published Online: September 2009

I. INTRODUCTION

The pricing and production of audit services has been the subject of much research since the seminal work of Simunic (1980). One specific line of research has examined the production process that comprises the audit (e.g., O'Keefe et al. 1994; Stein et al. 1994; Hackenbrack and Knechel 1997). O'Keefe et al. (1994) presented a theoretical model of audit production that treated auditor effort (e.g., hours) as input and the level of assurance obtained as output. Since the output of the audit (achieved assurance) is not observable, and limited data have been available to analyze audit production, it has been very difficult to empirically test the theoretical model directly. As a result, researchers have had to adapt theoretical models to fit the available data. For example, disaggregated labor hours for an engagement and client-specific characteristics have been used to measure, in turn, both the input and the output of an audit (O'Keefe et al. 1994; Hackenbrack and Knechel 1997; Dopuch et al. 2003).

We develop a modified audit production framework that uses an alternative output measure to proxy for the amount of evidence gathered during the course of the engagement. This approach allows us to directly derive an empirical model of audit production from the theoretical model developed by O'Keefe et al. (1994). We define the cost of labor as input to the auditor's production process. More labor reflects more input, so cost is a monotonically rising measure of production input. We argue that the labor input produces observable outputs that proxy for a specific level of audit evidence that is related to the unobservable outcome of the audit (i.e., the achieved level of assurance). Specifically, our output measure is the time an auditor spends on individual evidence-gathering activities, e.g., time spent on substantive testing measured in hours. The use of this measure assumes that more evidence is better than less, i.e., the lower the residual audit risk and the higher the level of audit assurance achieved. We also assume that the extent of evidence-gathering activities varies in a systematic manner with variations in the unobservable level of achieved assurance. Client-specific characteristics are treated as exogenous factors and assumed to affect audit production as a whole.

We first apply our modified audit production framework to measure the efficiency of individual audits using Data Envelopment Analysis (DEA; Cooper et al. 2007). We find that the appropriate form of DEA incorporates restrictions on non-critical evidence-gathering activities. We also illustrate an empirical technique for evaluating the relative trade-off among alternative procedures when some activities are deemed to be more important than others. Overall, we observe that only 15 percent of the audits in our sample were performed efficiently. Furthermore, there is substantial variation in efficiency across firm offices, and in some offices inefficiency is associated with higher fees.

Next, we examine the effect of exogenous client and engagement characteristics on audit efficiency using regression analysis. Our second-stage regression results suggest that the audits of larger clients, clients with extensive system automation, and clients with a December fiscal year-end are generally more efficient, while audits of clients that receive tax services from the auditor, clients that have subsidiaries, and engagements where the auditor relies on internal control are less efficient. Finally, we find that a well-specified traditional regression model of audit production controls for many of the circumstances that influence overall audit efficiency, and that audit report lag, while being a reasonable proxy for abnormal auditor effort, is not a proxy for auditor efficiency per se.

There are five major contributions of this research: First, it addresses a significant conundrum that has been present in virtually all research on audit production; namely, how to translate the theoretical model that links effort (inputs) to achieved assurance (output) into a viable empirical model. We do this by developing an intuitively appealing, and observable, output proxy that can vary with effort but is monotonically linked to the unobservable level of assurance. This allows us to remove the critical assumption underlying most prior research that all audits conducted by a firm within a given time period achieve a constant level of assurance. Second, our model separates the client structure from the audit production process so that we can examine these effects separately. Third, we develop and test a technique for evaluating potential trade-offs among different procedures that comprise the audit process. Fourth, we find that the residuals from standard regression models of auditor effort are highly correlated with audit efficiency. Fifth, we determine that certain segments of audit markets based on industry and location may not be competitive since they reveal an association between unexplained audit fees and audit efficiency.

Section II discusses a production-based approach to the audit process, the conventional audit production function, relevant prior research, and our modified audit production model. Section III describes the empirical model, our data, and descriptive statistics; Section IV presents our primary analysis and results; and Section V concludes with a summary of our results in the context of previous research.

II. MODELING AUDIT PRODUCTION

In most organizations, economic activities are organized by production units that convert inputs into outputs using processes that comprise a particular production technology. Following Shepherd (1970, 4), a *production function* is a mathematical expression of the relationship between the factors of production (inputs) and the outputs of the process. For an audit, the inputs to the audit process include the units of labor necessary to produce a given level of output. The output of an audit is more difficult to define. O'Keefe et al. (1994, 241) define the output of the audit production process as the achieved (but unobservable) level of assurance that financial statements are not materially misstated.

In early audit research, a lack of data led researchers to base inferences about audit production on audit fees. Simunic (1980) found that factors related to auditee size, complexity, asset composition, industry, risk, and financial distress, as well as auditor tenure and auditor type, are associated with audit fees. Simunic's (1980) study spawned a large stream of related research (see Hay et al. [2006] for a summary). However, evidence based on audit fees permits only indirect tests of audit production since an audit fee is the product of quantity (audit effort exerted by various grades of labor) and price (charge rate per unit of labor). Therefore, a change in audit fees may not reflect pure changes in audit production but merely changes in pricing. As a result, tests based on audit fees may be confounded by an audit firm's pricing policies (O'Keefe et al. 1994) and competitive market conditions (Hay and Knechel 2009).

A limited number of studies have used information on audit effort to test hypotheses about the production of audit services directly. As noted by O'Keefe et al. (1994), the output of audit production is the level of assurance achieved, while input measured by labor hours disaggregated by rank (partner, manager, supervisor, and staff) is the effort required to produce that output. This specification of the audit production process is based on three key assumptions: (1) the level of assurance produced by an audit firm is not observable, does not vary across engagements, and is implied by the audit firm's brand name; (2) the firm selects an auditor that delivers the desired level of assurance; and (3) the audit firm operates in a competitive environment. Based on these assumptions, O'Keefe et al. (1994) describe the auditor's decision problem as a constrained cost-minimization problem for a fixed level of assurance (see also Hackenbrack and Knechel 1997, 485). A potential audit supplier for a client is characterized by the desired assurance level delivered at a minimum cost. The resulting empirical model estimates actual audit hours as a function of client and engagement characteristics.

Early production studies use regression analysis and specify total audit hours spent on an engagement as the dependent variable, while client and engagement characteristics are included as independent variables (see Palmrose 1986, 1989; Davis et al. 1993; Davidson and Gist 1996). A number of other production studies extend this research by disaggregating audit hours by staff level (O'Keefe et al. 1994; Stein et al. 1994), or by staff level and audit activity (Hackenbrack and Knechel 1997). The use of disaggregated hours spent on engagements is appropriate when client and engagement factors have differential effects on the various types of labor. O'Keefe et al. (1994) find that specific client size and risk measures significantly influence the mix of labor resources: (1) increased client size results in greater use of audit staff, (2) increased client financial risk results in greater use of audit partners and managers, and (3) increased levels of inherent risk result in greater use of audit seniors and staff. Hackenbrack and Knechel (1997) further disaggregate labor hours by the type of audit activities performed, e.g., audit planning, substantive testing, and internal control evaluation, and find that client-specific characteristics have a significant impact on task assignment patterns and resource allocation decisions. By disaggregating labor hours by rank and audit activity, Hackenbrack and Knechel (1997) provide the foundation for the modified audit production model developed in this study.

Dopuch et al. (2003) extend the audit production literature by examining relative efficiency in audit production using efficiency frontier techniques such as Data Envelopment Analysis (DEA) and Stochastic Frontier Analysis (SFA). The unit of analysis for SFA and DEA is a decision-making unit (DMU), e.g., a firm, department, branch, or, in this study, an audit engagement. We use DEA (Cooper et al. 2007), which applies linear programming techniques to estimate an efficiency frontier formed by efficient DMUs that bound the inefficient ones. The shape of the efficiency frontier and how tightly it envelops the DMUs can vary depending on whether constant or variable returns to scale are assumed.[1] We discuss DEA in more detail further below and in Appendix A.

Dopuch et al. (2003) examine several empirical specifications of the audit production model. One specification treats client-specific characteristics as outputs and labor hours as inputs with the objective of maximizing the characteristics given the amount of labor hours used. In this approach, an efficient audit is interpreted as one that results in the highest level of client characteristics under audit (e.g., assets, subsidiaries, risk) given a certain level of labor. Conversely, and consistent with most prior audit production research, an alternative specification considers labor hours as outputs and client characteristics as inputs, with the objective of minimizing the labor hours given the client characteristics. In this case, an efficient audit is one that "produces" the lowest level of labor hours given the client characteristics. We take a different approach by assuming the effect of client characteristics to be exogenous, indirect, and pervasive to audit production process as a whole.

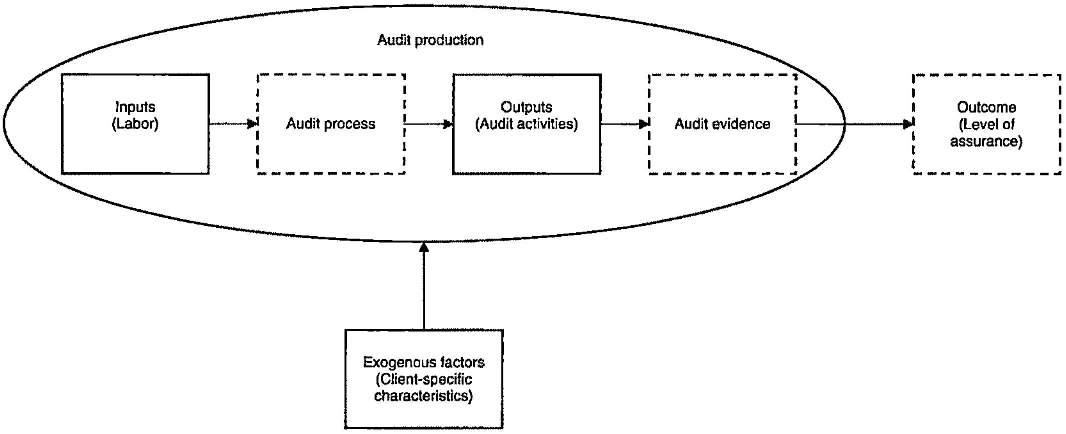
As previously discussed, the theoretical literature on audit production and efficiency identifies labor hours as inputs and the level of assurance (representing audit quality) as the output of the audit production process. This study builds on the original theoretical model developed in O'Keefe et al. (1994) by focusing on how increased labor converts to increased assurance using a different type of output measure than available in previous research. The two key aspects of the model are depicted in Figure 1. First, the model assumes labor inputs are used in the audit process to perform audit activities that result in the generation of audit evidence. There are different types of audit activities that may be performed, but they all lead to the generation of audit evidence. Thus, the greater the extent of evidence-gathering activities, the greater the likelihood that an auditor reaches the correct conclusion about whether the financial reports are misstated, leading to a higher achieved level of assurance (i.e., the "final" outcome of the audit). Second, client-specific characteristics are considered to be exogenous to the audit production model, affecting inputs, process, evidence-gathering activities (outputs), and evidence as a whole. Thus, they are not treated as either inputs or outputs, but they are assumed to have an effect on audit production and on the relative efficiency of an engagement.

---

[1] Constant returns to scale (CRS) assumes that a proportional change in inputs results in a corresponding proportional change in outputs.

This modified framework is based on Hackenbrack and Knechel (1997) who disaggregated labor hours into a matrix, classifying the hours by staff level and by type of audit activity. Hackenbrack and Knechel (1997) showed that client-specific characteristics affect the matrix of labor hours as a whole, suggesting that these characteristics are exogenous to the audit production process. If client-specific characteristics are treated as exogenous factors, then it is likely that they affect both the use of different levels of audit staff *and* the extent of specific activities performed by audit staff. While the level of assurance can be seen as the ultimate outcome of an audit, audit activities that result in the generation of audit evidence are regarded as "intermediate" outputs. Specifically, we assume that increased hours spent on activities such as audit planning, substantive testing, and internal control evaluation reduce the probability that the financial statements are materially misstated and thus serve as a reasonable proxy for the unobservable level of achieved assurance. Furthermore, we assume the evidence-gathering activities are substitutable since they are all performed to increase the probability that the required level of assurance will be achieved.[2]

FIGURE 1
A Modified Audit Production Framework


Audit production
Inputs (Labor)
Audit process
Outputs (Audit activities)
Audit evidence
Outcome (Level of assurance)
Exogenous factors (Client-specific characteristics)

---

[2] While the activities are probably not perfectly substitutable—no audit will consist of only planning or internal control testing—the audit risk model presumes that different evidence-gathering activities are substitutable.

We use DEA to determine relative audit production efficiency. DEA requires no assumptions as to the functional form of the production models, i.e., how inputs are converted into outputs, making it particularly suitable for our setting. DEA can measure efficiency using either an input or output orientation. With an input orientation, a DMU is considered efficient if its inputs cannot be proportionally reduced without changing the output produced, i.e., the amount of input is minimized for a given level of output. With an output orientation, a DMU is assumed efficient if its output cannot be increased without changing the inputs, i.e., the amount of output is maximized for a given level of input. Since audit production is generally characterized as a minimization problem, we use an input orientation in our DEA application.

The selection of inputs and outputs is an important step in applying DEA. In general, outputs are assumed to increase with increases in input, and more output is generally considered better than less, while less input is generally considered better than more. However, there may be some outputs that are less desirable than others (Seiford and Zhu 2002; Färe and Grosskopf 2004) or have less "value." For example, pollution is typically regarded as an undesirable output that is generated as a by-product of production. For a factory to improve, its pollutants should be reduced or kept to a minimum. Similarly, some outputs, while having positive value, may be less valuable than the primary output, e.g., sawdust produced during the process of making lumber. In much the same way, not all activities performed in the audit process will contribute equally to an increase in achieved assurance (e.g., non-critical tests, engagement administration) and it may be desirable to limit or minimize these activities as they may be considered non-value adding. Our analysis specifically examines the potential trade-offs that can be made among different types of evidence-gathering activities while still achieving efficiency within an audit.

III. MODEL, DATA, AND DESCRIPTIVE STATISTICS

Basic Model

In our DEA application, the cost of staff resources is considered as the input and the disaggregated labor hours devoted to various evidence-gathering audit activities as the output (see the full model in Appendix B). We define our output measure using the eight activity categories defined in Hackenbrack and Knechel (1997): (1) audit planning, (2) internal control evaluation, (3) substantive testing—critical objectives, (4) substantive testing—non-critical objectives, (5) review—critical, (6) review—non-critical, (7) financial statement preparation, and (8) client interaction. Note that an important assumption underlying this approach is that these activities yield audit evidence, so "more is better than less," i.e., *ceteris paribus*, the more time spent on these activities, the more evidence the auditor will have and the higher will be the achieved (but unobservable) level of overall assurance. In other words, the marginal effect of another hour of audit activity devoted to gathering evidence is strictly positive.[3]

The input for our DEA model reflects the total cost of personnel. The more input in the form of hours charged to an engagement, the higher the cost of the audit. The overall labor costs are derived by multiplying labor hours by hourly charge-out rates for different levels of labor: partner, manager, supervisor, and staff. While charge-out rates are not a perfect proxy for costs as they include a profit margin, they reflect the differential costs of the different levels of labor. As previously noted, client attributes and characteristics are assumed to be exogenous and are not considered as either inputs or outputs in our framework. Rather, these characteristics are presumed to influence the relationship between inputs (labor cost) and outputs (hours spent on evidence-gathering activities). We use regression to analyze the effect of client and engagement characteristics on the efficiency scores across engagements in the second stage of our analysis.

---

[3] The issue of marginal *value* of audit evidence is important but not specifically addressed in this study. The marginal value of another hour of evidence-gathering (presumably due to the reduced risk of issuing the wrong opinion) may or may not always exceed the marginal cost of obtaining that evidence.

Evaluating Trade-Offs among Evidence-Gathering Procedures

We further refine our output measures by considering the relative value of the different evidence-gathering activities. To illustrate, we treat six of the intermediate activities as "high-value" outputs and two as "low-value." The two low-value evidence-gathering activities are: (4) Substantive—Non-Critical and (6) Review—Non-Critical. This distinction is made based on the assumption that these outputs do not contribute as much to overall assurance as the other activities, and also reflect the official audit policies of the firm providing the data. To validate the determination of non-critical activities, we reviewed a sample of engagements with audit partners within the firm to determine if "non-critical" work was considered to be necessary on a specific engagement. The partners generally concluded that "non-critical" work was significantly less important for audit quality and much of this work could have been omitted from the audit without compromising audit effectiveness. Consequently, while these low-value outputs may be an inevitable component of any audit (e.g., the audit equivalent of sawdust), an audit should not be considered "efficient" if a substantial amount of these activities are produced.

Based on the distinctions in evidence-gathering activities, we conducted a number of different analyses. First, we separately analyzed the six "good" outputs, omitting the two low-value outputs entirely. Second, still omitting the low-value outputs, we assume that the amount of time spent on reviewing critical substantive tests should not exceed the time spent conducting the tests, so a constraint is imposed that the ratio of Substantive Testing—Critical to Review—Critical should be more than one for an audit to be considered to be efficient.[4] Next, we analyzed all eight outputs. Simply including two more evidence-gathering activities increases the efficiency of all audits[5] unless constraints are imposed on the value of non-critical activities relative to the other activities. To incorporate such constraints, we adopt the following restrictions (see Appendix B for more detail):

- The ratio of Substantive Testing—Critical to Substantive Testing—Non-Critical should be equal to or greater than s.
- The ratio of Review—Critical to Review—Non-Critical should be equal to or greater than s.

These constraints specify the proportional value of comparable critical and non-critical activities. The scalar s reflects the relative trade-off in value of the paired activities. For example, if s equals 2, then critical testing is considered to be twice as valuable as non-critical testing. The greater the value of s, the less valuable non-critical activities are considered to be.

---

[4] To be more specific, we define y_r as the rth output (hours spent on evidence-gathering activity r) and μ_r is the weight placed on the rth activity ($r = 1,\ldots,8$). The constraint on the review of critical substantive tests implies that $\mu_3 \geq \mu_5$, where $r = 3$ indicates the activity Substantive testing—Critical and $r = 5$ indicates the activity Review—Critical (see Appendix B for more detail).

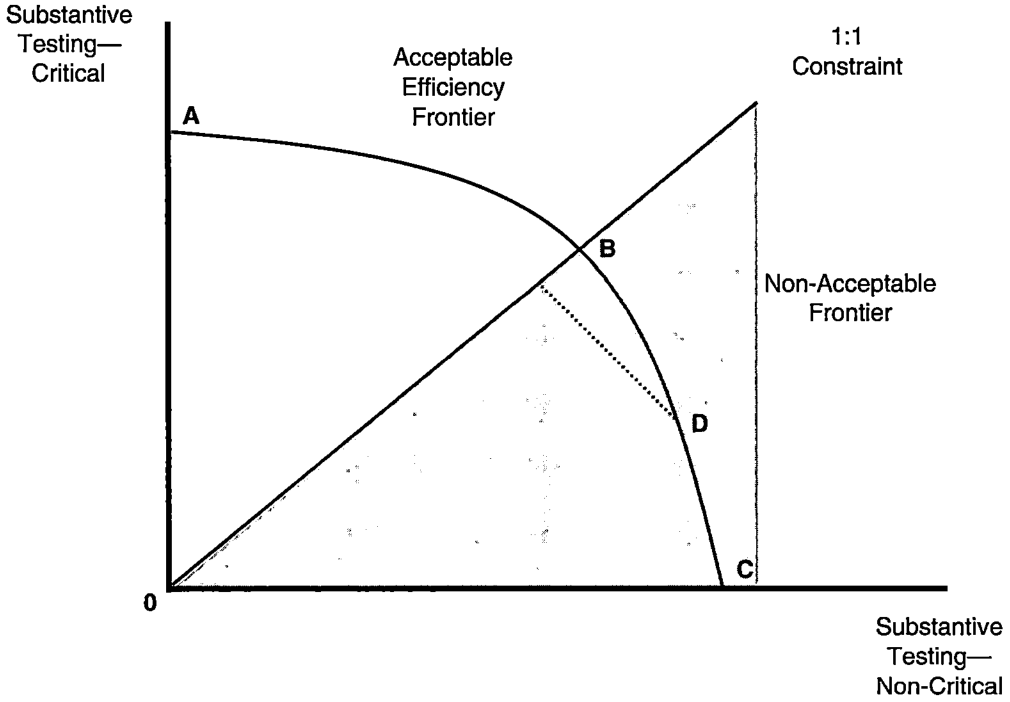
[5] The number of outputs indicates the dimensionality of the efficiency frontier. In eight dimensions (eight outputs), all audits are no farther from the frontier—and may in fact be much closer—than they would be in six dimensions.

To illustrate the effect of these constraints, consider Figure 2 depicting a two-dimensional mapping of the relationship between Substantive Testing—Critical and Substantive Testing—Non-Critical. Given that more outputs are better than less, the efficiency frontier is convex to the origin. In an unconstrained environment, any combination of outputs on the ABC arc would be efficient for a given level of input. However, our model limits the amount of non-critical testing in an efficient audit, effectively constraining the frontier with a ray that has slope s and originates at the origin. In the example, we truncate the frontier at the point where $s = 1$, i.e., hours devoted to non-critical testing cannot exceed critical testing. Consequently, the efficient frontier is reduced to the arc AB. Any audits falling on the arc BC would be deemed inefficient, i.e., point D would be assigned an efficiency score of 1.00 in an *unconstrained* environment but is considered to be inefficient in a *constrained* environment. The effect of the weight constraints based on s is to decrease the efficiency of audits with more time devoted to non-critical activities. The value of s where an auditor would be indifferent between critical and non-critical activities can be determined using tests developed by Banker (1993) and Banker and Natarajan (2004).

Overall, we examine six DEA models that differ in the number of outputs that are considered and the types of constraints that are imposed:

- Model (1): Six high-value outputs with no restrictions.
- Model (2): Six high-value outputs with restriction that Review—Critical cannot exceed Substantive—Critical.

FIGURE 2
The Effect of Constraints on the Trade-Off of Selected Outputs


Substantive Testing—Critical
Acceptable Efficiency Frontier
1:1 Constraint
A
B
Non-Acceptable Frontier
D
C
0
Substantive Testing—Non-Critical

- Model (3): Eight outputs without restrictions (i.e., non-critical activities considered equally valuable as all other activities).
- Model (4): Eight outputs with restriction that Review—Critical cannot exceed Substantive—Critical.
- Model (5): Eight outputs with restriction that Review—Critical cannot exceed Substantive—Critical *and* restriction of $s = 8$ for non-critical activities.
- Model (6): Eight outputs with restriction that Review—Critical cannot exceed Substantive—Critical *and* restriction of $s = 11$ for non-critical activities.

Models (5) and (6) reflect differences in the test statistics used to derive s (Banker 1993). We analyze all six models using both constant returns to scale (CRS) and variable returns to scale (VRS), resulting in 12 unique sets of efficiency results.

Data

To apply our model, we use the same data set as in Hackenbrack and Knechel (1997) and Knechel and Payne (2001). The data was obtained by means of a survey developed and administered by a large international audit firm relating to U.S.-based audits with a 1991 year-end.[6] The population was stratified into nine industry and two size categories. For each stratum, 25 engagements were selected, resulting in a total sample size of 450 engagements. The surveys were sent to the respective engagement partners and completed by the engagement team. The survey was used to create a data set containing the number of labor hours by rank charged to various audit activities, as well as various client and engagement characteristics. Of the 327 responses received, 307 are used to compute the efficiency scores using DEA. Seventeen responses were eliminated because they were for non-audit engagements, one response had insufficient data, and two were dropped as outliers based on their super-efficiency scores (see Appendix A). Of the 307 responses used for the DEA analysis, 219 had adequate data to be included in the second-stage regression analysis.

Descriptive Results

Panel A of Table 1 reports the descriptive statistics on labor costs and hours by audit activities. The mean disaggregated activity hours suggest that substantive testing in both the critical (500) and non-critical (269) areas consume the largest amount of engagement effort. Although the firm supposedly relied on internal control in many audits (61 percent, Panel B), the actual amount of time spent evaluating internal control was relatively small (53 hours). The wide range of values observed for the various audit activities supports our assumption that audit activities are substitutes in "producing" the required level of assurance. All input and output variables are highly skewed. Panel B contains the exogenous factors used in the second-stage regression and will be discussed later under the multivariate tests.

IV. ANALYSES AND RESULTS

Based on the theoretical specification of our modified audit production framework we use an input-orientation DEA model. We treat evidence-gathering activities (outputs) as fixed and assume labor is varied to minimize costs. We use both constant (CRS) and variable (VRS) returns to scale models in our analyses and report the results in Table 2. A CRS

---

[6] A key advantage of using cross-sectional data from a single audit firm is that it reduces the confounding effects of pricing policies and differences in production technologies (O'Keefe et al. 1994, 243).

TABLE 1
Descriptive Statistics for Variables Used in the DEA and Regression Models

Panel A: Audit Production Data (n = 307)[a]

	Mean		Std. Dev.	Median	Min.	Max.
DEA Input – Input Costs						
Staff Costs ($)	141,350		164,097	89,958	7,999	1,395,507
DEA Output – Activity Hours						
Audit Planning	106	(7%)	158	60	3	1,280
Internal Control	53	(4%)	80	25	0	583
Substantive Testing—Critical	500	(35%)	630	268	0	5,470
Substantive Testing—Non-Critical	269	(19%)	288	176	0	1,708
Review—Critical	146	(10%)	184	79	0	1,100
Review—Non-Critical	65	(5%)	94	33	0	850
Financial Statements	131	(9%)	240	62	0	3,243
Client Interaction	153	(11%)	267	78	0	2,590
Total Activity Hours[b]	1,421	(100%)	1,599	950	88	14,147

Panel B: Client and Engagement Characteristics (n = 226)

	Mean	Std. Dev.	Median	Min.	Max.
CATREV	7.63	2.42	8	1	11
NEWCLNT	0.23	0.42	0	0	1
TAXYN	0.77	0.42	1	0	1
MASYN	0.39	0.49	0	0	1
DELAY	71.71	50.79	59	10	442
GEODISP	0.38	0.49	0	0	1
ORGCOMPL	0.42	0.50	0	0	1
SUBYN	0.58	0.49	1	0	1
TIMING	0.38	0.49	0	0	1
AUTOM	0.54	0.50	1	0	1
ICRELY	0.61	0.49	1	0	1
FEERESID	−0.00	0.69	−0.03	−3.61	2.21
PUBLIC	0.26	0.44	0	0	1
YEAREND	0.58	0.50	1	0	1

---

[a] Input is total labor costs (in $) and outputs are disaggregated labor hours spent on various activities, stratified into eight categories as in Hackenbrack and Knechel (1997).

[b] The difference between the addition of average hours per activity and total activity hours is due to rounding.

Variable Definitions:

CATREV = size of the client measured by revenues, classified into 11 categories;
NEWCLNT = 1 if the duration of auditor tenure is three years or less, 0 otherwise;
TAXYN = 1 if tax services are provided to the client, 0 otherwise;
MASYN = 1 if consulting services are provided to the client, 0 otherwise;
DELAY = number of days between financial year-end and the date of the audit report, a measure of the extent of audit report lag;
GEODISP = 1 if the client is geographically dispersed, 0 otherwise;
ORGCOMPL = 1 if the client's organizational structure as assessed by the audit team is complex, 0 otherwise;
SUBYN = 1 if the client has subsidiaries, 0 otherwise;
TIMING = 1 if extensive interim work is performed, 0 otherwise;
AUTOM = 1 if the client is extensively automated, 0 otherwise;

(*continued on next page*)

TABLE 1 (continued)

ICRELY = 1 if the audit team indicates high reliance on the client's internal controls, 0 otherwise;
FEERESID = residual from an audit fee model estimated using the companies in the sample;
PUBLIC = 1 if the client is publicly listed, 0 otherwise; and
YEAREND = 1 if the client has a December year-end, 0 otherwise.

TABLE 2
DEA Results for 307 Audit Engagements Using Different Assumptions about Outputs[a]

	Models Using Six Outputs		Models Using Eight Outputs			
	Model (1) No Weight Restrictions	**Model (2)** Weight Restriction on Review—Critical	**Model (3)** No Weight Restrictions	**Model (4)** Weight Restriction on Review—Critical	**Model (5)** Weight Restrictions on Non-Critical Activities (s = 8)	**Model (6)** Weight Restrictions on Non-Critical Activities (s = 11)
Constant Returns to Scale (CRS)						
Mean	82.11%	80.53%	93.85%	92.59%	82.06%	81.64%
Median	82.80%	80.28%	94.41%	92.89%	82.03%	81.39%
Std. Dev.	12.72%	12.81%	5.97%	6.50%	11.79%	12.08%
Min.	23.96%	23.96%	71.08%	69.08%	25.48%	24.36%
Nr. (%) efficient	42 (13.7%)	37 (12.1%)	85 (27.7%)	76 (24.8%)	37 (12.1%)	37 (12.1%)
Variable Returns to Scale (VRS)						
Mean	84.99%	83.01%	95.58%	94.36%	84.45%	84.05%
Median	86.00%	82.74%	98.02%	95.98%	83.92%	83.51%
Std. Dev.	12.84%	12.96%	5.47%	6.15%	11.90%	12.19%
Min.	33.52%	33.52%	76.82%	70.01%	33.52%	33.52%
Nr. (%) efficient	68 (22.1%)	48 (15.6%)	133 (43.3%)	111 (36.1%)	48 (15.6%)	48 (15.6%)

---

[a] Notes on different models:
Model (1): Six high-value outputs with no weight restrictions;
Model (2): Six high-value outputs with restriction that Review—Critical cannot exceed Substantive—Critical;
Model (3): Eight outputs without restrictions (i.e., non-critical activities considered equally valuable as all other activities);
Model (4): Eight outputs with restriction that Review—Critical cannot exceed Substantive—Critical;
Model (5): Eight outputs with restriction that Review—Critical cannot exceed Substantive—Critical *and* weight restriction that s = 8 (derived based on the exponential distributional assumption, see Banker [1993] for details of test statistic); and
Model (6): Eight outputs with restriction that Review—Critical cannot exceed Substantive—Critical *and* weight restriction that s = 11 (derived based on the half-normal distributional assumption, see Banker [1993] for details of test statistic).

model assumes that changes in inputs are strictly proportional to changes in output, while a VRS model assumes that as inputs increase, outputs increase at a different rate (e.g., doubling input does not yield double the output). As discussed in more detail below, CRS and VRS yield significantly different results in some of our analyses, suggesting that the size of an engagement as measured by total labor cost (e.g., the "scale") may differentially affect efficiency.

Data Envelopment Analysis: Efficiency Results for Individual Engagements

DEA Analysis Based on the Six-Output Model: Models (1) and (2)

Models (1) and (2) in Table 2 serve as our baseline and reflect the results based on an analysis of six outputs, omitting the activities classified as non-critical. For Model (1), the model with no weight restrictions, VRS yields an average efficiency of 84.99 percent, and 22.1 percent of the engagements are considered to be efficient. Efficiency is lower using a CRS model: average efficiency is 82.11 percent, and 13.7 percent of the engagements are considered to be efficient. These results suggest engagement size may affect efficiency.[7] Based on this result, and given that size has been demonstrated to have a significant effect in models of audit fees and audit effort, we conclude that the VRS approach is appropriate for the analysis of audit production and focus on those results. We next add a weight restriction on reviewing critical substantive tests in Model (2). This restriction is based on the assumption that no more time should be spent reviewing substantive tests than was spent performing them. The effect of this restriction is to reduce average efficiency to 83.01 percent[8] and the proportion of efficient audits to 15.6 percent.

DEA Analysis Based on Eight-Output Model: Models (3) through (6)

The effect of increasing outputs from six to eight has a noticeable effect on audit efficiency. In Model (3) with no weight restrictions, adding two outputs increases efficiency scores by approximately 11 percent compared to Model (1) (84.99 percent to 95.58 percent). The proportion of efficient audits increases from 22.1 percent to 43.3 percent. The imposition of the constraint on review of critical substantive testing in Model (4) has a small, but significant, impact on the mean efficiency scores (94.36 percent for Model (4) compared to 95.58 percent for Model (3)), and the proportion of efficient audits drops from 43.3 percent to 36.1 percent.

We next consider restrictions on the non-critical activities in Models (5) and (6). We derive the scalar s that reflects the point of indifference between critical and non-critical audit activities by incrementally increasing its value until we obtain efficiency scores in the constrained eight-output model that are statistically equivalent to the six-output model. Note that this analysis always incorporates the restriction on the review of critical substantive testing, i.e., we start with Model (4). Three different tests of statistical equivalence are used (Banker and Natarajan 2004): exponential, half-normal, and Kolmogorov-Smirnov. Using a significance level of 0.10, the exponential and Kolmogorov-Smirnov tests indicate that s is approximately equal to 8. For the half-normal test, s is approximately 11. We conclude that between 8 to 11 hours of low-value activities (Substantive Testing—Non-Critical and Review—Non-Critical) are equivalent to one hour spent on comparable high-value activities. This does not imply that these activities are necessarily bad (e.g., pollution), but they provide less assurance than other evidence-gathering activities.

Models (5) and (6) exhibit a significant decrease in efficiency when compared to the other eight-output models, with a decrease in average efficiency from 94.36 percent (Model 4) to 84.45 percent (Model 5) and 84.05 percent (Model 6). Also, the number of efficient engagements drops from 36.1 percent to 15.6 percent in both constrained models. We further note that the average efficiency when $s = 11$ is not significantly lower than the average efficiency when $s = 8$, nor do the number of efficient engagements differ. Consequently, we conclude that the marginal effect of increasing s from 8 to 11 is minimal and focus our subsequent analysis on Model (5). Taken together, these results suggest a need for firms to identify and reduce low-value staff activities, while also recognizing that some level of low-value activity is a natural part of any audit.[9]

---

[7] The differences between the VRS and CRS results are statistically significant (Banker and Natarajan 2004).

[8] All differences in efficiency results between models that are reported are statistically significant with one exception: CRS results for Models (1) and (2) do not differ significantly from each other.

DEA Results Considering Environmental Factors

The fact that applying VRS affects our efficiency scores suggests that there may be other environmental factors that could influence audit efficiency. Following Fried et al. (2002) and Banker and Natarajan (2004), we investigate the effects of a number of these factors by constructing a second-stage DEA model where the inefficiency results from Model (5) are treated as inputs, and specific environmental factors are treated as non-controllable outputs. We consider four client factors as outputs of this second-stage analysis: number of subsidiaries, geographical dispersion, organizational complexity, and client size. The assumption is that higher values of these environmental factors may be associated with a loss of efficiency since auditors will generally need to spend more time on audits of large, dispersed and complex clients. For an inefficient audit, the second-stage DEA allows us to separate environmental inefficiency (considered inherent in the engagement) from managerial inefficiency (considered controllable by the auditor).[10] We did not find a significant amount of environmental inefficiency for any of the four measures of client size or complexity, either individually or in combination. This result was the same for CRS and VRS.[11] We conclude that a large proportion of the inefficiency observed in our basic analysis is attributable to managerial inefficiency.

Efficiency across Offices and Industries

Audit firms are generally organized by office, and professionals often specialize in specific industries. Thus, engagements can be aggregated into clusters based on those two dimensions. Table 3 reports the results of our efficiency analysis for Model (5) by office and industry. We only report results for offices that have at least ten audits in our sample.[12] The average efficiency across industries ranges from a low of 80 percent (high tech) to a high of 86 percent (financial institutions). A greater variance can be observed in the number of audits that are deemed to be efficient, with only 5.00 percent, 6.25 percent, or 7.41 percent of audits in high tech, merchandising, and health care, respectively, being efficient, compared to 24.24 percent of governmental audits. Since efficiency scores are computed relative to the entire sample, these results suggest that it may be misleading to ignore industry when comparing audit efficiency.

Efficiency across offices ranges from a low of 79 percent (office 15) to a high of 89 percent (offices 105, 147, 181, and 293). The proportion of efficient audits ranges from a low of 4.76 percent (office 15 again) to a high of 28.57 percent (office 166). Since industry specialization may vary by office (Ferguson et al. 2003; Francis et al. 2006), the variations in efficiency across offices may be related to the differences in efficiency across industry.

---

[9] For example, low-value audit activities, especially those assigned to junior staff, may be highly valuable to the firm as a whole if they represent significant training opportunities for the staff. Such benefits, however, are external to a specific engagement.

[10] An engagement that was found efficient in the first stage would continue to be rated as efficient in the second-stage analysis.

[11] Note, the use of VRS in the second stage of the analysis may actually overcompensate for environmental factors since the size and complexity measures we use in the analysis are highly correlated with the extent of inputs used in the first stage.

[12] The firm providing the data considered information about the specific location of offices to be sensitive and proprietary.

TABLE 3
DEA Results Classified by Industry and Office Based on Model (5)

	n	Mean	Std. Dev.	Median	Min.	Max.	Nr. (%) of Efficient DMUs
Sample by Industry							
Manufacturing	48	0.84	0.12	0.83	0.57	1	9 (18.75%)
Finance, Insurance and Real Estate	115	0.86	0.11	0.86	0.53	1	21 (18.26%)
Government	33	0.85	0.14	0.85	0.34	1	8 (24.24%)
Health Care	27	0.83	0.13	0.83	0.57	1	2 (7.41%)
High Tech	20	0.80	0.14	0.82	0.47	1	1 (5.00%)
Merchandising	32	0.83	0.10	0.81	0.62	1	2 (6.25%)
Education	32	0.85	0.11	0.83	0.68	1	5 (15.63%)
Sample by Office[a]							
Office 15	21	0.79	0.12	0.81	0.47	1	1 (4.76%)
Office 69	21	0.81	0.15	0.80	0.53	1	5 (23.81%)
Office 75	13	0.80	0.15	0.80	0.54	1	2 (15.38%)
Office 105	11	0.89	0.11	0.92	0.71	1	3 (27.27%)
Office 115	13	0.86	0.10	0.84	0.71	1	1 (7.69%)
Office 147	17	0.89	0.11	0.95	0.71	1	4 (23.53%)
Office 163	10	0.85	0.10	0.85	0.70	1	2 (20.00%)
Office 166	14	0.88	0.11	0.86	0.69	1	4 (28.57%)
Office 181	11	0.89	0.09	0.92	0.77	1	2 (18.18%)
Office 182	12	0.82	0.11	0.82	0.64	1	1 (8.33%)
Office 194	13	0.83	0.09	0.84	0.68	1	1 (7.69%)
Office 196	20	0.86	0.10	0.83	0.68	1	3 (15.00%)
Office 204	12	0.86	0.09	0.85	0.67	1	2 (16.67%)
Office 221	13	0.85	0.09	0.86	0.71	1	1 (7.69%)
Office 251	19	0.86	0.13	0.88	0.60	1	3 (15.79%)
Office 293	15	0.89	0.12	0.91	0.57	1	4 (26.67%)

[a] Only efficiencies for offices with ten or more engagements are reported. All industry categories have ten or more engagements and therefore efficiencies for all industry categories are reported.

Another possible explanation is that different offices have different levels of skilled personnel. This may be due to the size of the office as recently suggested by Francis and Yu (2009), or it may be due to other factors that are codependent with office location such as distance from the office to a client (Jensen et al. 2008).

Summary of DEA Results

Based on the above results, we conclude that the most appropriate model for our sample is the VRS version of Model (5) with $s = 8$ (average efficiency of 84.45 percent, 15.6 percent rate of efficient engagements). These results are comparable to Dopuch et al. (2003) who report an overall efficiency of 88.3 percent, with 14.2 percent of the audits on the efficiency frontier. For audits whose efficiency score is less than 100 percent, average efficiency is 81.6 percent (not reported), suggesting that the inputs on these engagements could be reduced by over 18 percent. The relatively low level of structural efficiency might be a surprise given that large audit firms have clearly stated formal policies regarding their standard of work and staffing requirements, i.e., similar to the concept of a professional culture described in Rouse et al. (1997). However, it is clear from Table 2 that when weight restrictions are imposed, a large number of audits lose efficiency due to excess effort for non-critical audit tests, e.g., possible over-auditing of low-risk areas.

Regression Analysis: Impact of Exogenous Factors

The Second-Stage Regression Model

In our audit production framework, client and engagement characteristics are considered exogenous to audit production. We test this by examining how client-specific characteristics affect the efficiency scores of individual audits. We estimate the following regression model based on results from prior research:[13]

$$\begin{aligned} LNEFF = {} & \beta_0 + \beta_1 CATREV + \beta_2 NEWCLNT + \beta_3 TAXYN + \beta_4 MASYN \\ & + \beta_5 DELAY + \beta_6 GEODISP + \beta_7 ORGCOMPL + \beta_8 SUBYN \\ & + \beta_9 TIMING + \beta_{10} AUTOM + \beta_{11} ICRELY \\ & + \beta_{12} FEERESID + \beta_{13} PUBLIC + \beta_{14} YEAREND \\ & + \beta_{15}\text{–}\beta_{20} INDUSTRY + \beta_{21}\text{–}\beta_{48} OFFICE + \varepsilon \qquad (1) \end{aligned}$$

where:

LNEFF = natural logarithm of the efficiency score from Model (5) using VRS;[14]
CATREV = size of the client measured by revenues, classified into 11 categories;
NEWCLNT = 1 if the duration of auditor tenure is three years or less, 0 otherwise;
TAXYN = 1 if tax services are provided to the client, 0 otherwise;
MASYN = 1 if consulting services are provided to the client, 0 otherwise;
DELAY = number of days between financial year-end and the date of the audit report, a measure of the extent of audit report lag;
GEODISP = 1 if the client is geographically dispersed, 0 otherwise;
ORGCOMPL = 1 if the client's organizational structure as assessed by the audit team is complex, 0 otherwise;
SUBYN = 1 if the client has subsidiaries, 0 otherwise;
TIMING = 1 if extensive interim work is performed, 0 otherwise;
AUTOM = 1 if the client is extensively automated, 0 otherwise;
ICRELY = 1 if the audit team indicates high reliance on the client's internal controls, 0 otherwise;
FEERESID = residual from an audit fee model estimated using the companies in the sample. Audits that have fees higher or lower than expected may be more or less efficient as auditors adjust their work load to reflect the fee level;[15]
PUBLIC = 1 if the client is publicly listed, 0 otherwise;

---

[13] Determinants of (in)efficiencies have been investigated in a similar manner for other industries (Pitt and Lee 1981; Reifschneider and Stevenson 1991).

[14] We use the natural log of the efficiency scores (*LNEFF*) in our empirical model to control for heteroscedasticity.

[15] The fee regression model that we use is similar to the audit production model used in Hackenbrack and Knechel (1997). Specifically, the model estimates fees using independent variables for *CATREV*, *SUBYN*, *PUBLIC*, *ICRELY*, *MASYN*, and *FININD*. The model, estimated for 232 observations, has an overall adjusted R^2 of 0.661 and is significant at $p < 0.000$.

YEAREND = 1 if the client has a December year-end, 0 otherwise;
INDUSTRY = six dummy variables that represent the seven industries in the sample;[16] and
OFFICE = 28 dummy variables that represent the 29 offices that performed the audit engagements.

The effect of these variables on audit effort is well understood from prior studies, e.g., size and risk increase audit effort, while automation and years of experience reduce audit effort. However, since the efficiency scores used in our analyses reflect the efficiency, rather than the level, of effort in an engagement, the effect on specific audits is unknown so we do not hypothesize directional effects for the coefficients in the model.

Panel B of Table 1 reports descriptive statistics for the client characteristics. We observe that 23 percent of the firms have had their auditor for three years or less,[17] the firm provides tax services to its clients more frequently than consulting services (77 percent versus 39 percent), and it takes about 72 days to complete the audit. Less than half the companies in the sample are geographically dispersed (38 percent) or have a complex organizational structure (42 percent). A majority of the clients own subsidiaries (58 percent) and make extensive use of automated systems (54 percent). The firm performs relatively moderate interim work (38 percent of all audits), and claims to rely on the clients' internal control systems (61 percent of all clients). Finally, 26 percent of the companies are publicly traded and 58 percent have a December year-end. Correlations among the independent variables are reported in Table 4.

Determinants of Efficiency

Since the DEA efficiency scores are truncated at 1.00, we use a Tobit, or censored, regression model for our analysis (see Gujarati 1995, 570).[18] Our results are reported in Table 5. Despite the differences in efficiency scores between industries (see Table 3), industry classifications were not significant so Table 5 reports results excluding industry variables. Regression results indicated that some offices had significantly lower efficiencies than others so we include the office dummies that are significant.

Table 5 highlights three factors that are associated with an increase in audit efficiency in our sample: *CATREV*, *YEAREND*, and *AUTOM*. Specifically, audits of larger clients are generally more efficient. Since prior research has shown that the increase in audit effort for larger clients is disproportionately reflected in lower-cost staff accountants (O'Keefe et al. 1994), our finding suggests that this approach to staffing generally improves audit efficiency. Additionally, audits of clients with a December year-end are more efficient than other audits, possibly because auditors have less slack in staffing during the busy season so they are forced to be more efficient. Finally, audits are more efficient when a client has automated processes that may make it easier for auditors to retrieve and verify information relevant to the audit. Three factors are associated with reduced audit efficiency: *ICRELY*, *TAXYN*, and *SUBYN*. That is, audits are less efficient when the auditor relies on internal control, the client receives tax services from the audit firm, or the client has subsidiaries that are subject to audit. *SUBYN* is not controllable by the auditor and clients with subsidiaries may require more planning and coordination than is needed for the audit of companies without

---

[16] The original population was stratified into nine industry categories, but the reduced sample used for the regression analysis contains observations from only seven industry categories.

[17] The companies in the sample have had their current auditor for 11 years on average.

[18] Note, OLS yields very similar results. The OLS model has an adjusted R^2 of 0.237 and an F-statistic of 3.59 ($p < 0.000$). However, *SUBYN* is no longer significant, while *DELAY* is.

TABLE 4
Correlations among Independent Variables Included in Second-Stage Regression Model[a]

Variables	*CATREV*	*NEWCLNT*	*TAXYN*	*MASYN*	*DELAY*	*GEODISP*	*ORGCOMPL*	*SUBYN*	*TIMING*	*AUTOM*	*ICRELY*	*FEERESID*	*PUBLIC*
NEWCLNT	−0.129 (0.052)												
TAXYN	0.131 (0.049)	−0.020 (0.765)											
MASYN	0.455 (0.000)	−0.092 (0.170)	0.206 (0.002)										
DELAY	−0.094 (0.160)	0.058 (0.390)	−0.130 (0.051)	0.021 (0.755)									
GEODISP	0.239 (0.000)	−0.173 (0.009)	0.116 (0.082)	0.002 (0.972)	−0.115 (0.084)								
ORGCOMPL	0.508 (0.000)	−0.023 (0.729)	−0.010 (0.878)	0.287 (0.000)	−0.017 (0.802)	0.240 (0.000)							
SUBYN	0.376 (0.000)	−0.008 (0.906)	0.190 (0.004)	0.232 (0.000)	−0.219 (0.001)	0.262 (0.000)	0.344 (0.000)						
TIMING	0.428 (0.000)	0.064 (0.335)	0.116 (0.082)	0.301 (0.000)	−0.266 (0.000)	0.103 (0.123)	0.258 (0.000)	0.372 (0.000)					
AUTOM	0.467 (0.000)	−0.081 (0.225)	0.154 (0.020)	0.307 (0.000)	−0.171 (0.010)	0.117 (0.079)	0.316 (0.000)	0.294 (0.000)	0.318 (0.000)				
ICRELY	0.436 (0.000)	0.053 (0.425)	0.003 (0.967)	0.179 (0.007)	−0.124 (0.063)	0.135 (0.042)	0.216 (0.001)	0.239 (0.000)	0.340 (0.000)	0.321 (0.000)			
FEERESID	0.001 (0.986)	−0.054 (0.419)	−0.079 (0.239)	0.004 (0.951)	0.076 (0.258)	0.106 (0.112)	0.218 (0.001)	0.014 (0.830)	0.065 (0.331)	0.065 (0.334)	0.015 (0.828)		
PUBLIC	0.355 (0.000)	−0.014 (0.837)	0.210 (0.002)	0.187 (0.005)	−0.348 (0.000)	0.151 (0.023)	0.264 (0.000)	0.379 (0.000)	0.482 (0.000)	0.291 (0.000)	0.190 (0.004)	0.008 (0.905)	
YEAREND	0.079 (0.236)	0.023 (0.729)	0.158 (0.017)	0.044 (0.513)	−0.096 (0.151)	0.018 (0.793)	−0.022 (0.741)	0.237 (0.000)	0.349 (0.000)	0.223 (0.001)	0.114 (0.089)	0.060 (0.371)	0.225 (0.001)

(*continued on next page*)

TABLE 4 (continued)

---

[a] (Two-sided) p-values between parentheses. Reported coefficients are for Pearson product-moment correlations. Spearman rank correlations and phi correlations (as well as their significance levels) are virtually the same.

Variable Definitions:

CATREV = size of the client measured by revenues, classified into 11 categories;
NEWCLNT = 1 if the duration of auditor tenure is three years or less, 0 otherwise;
TAXYN = 1 if tax services are provided to the client, 0 otherwise;
MASYN = 1 if consulting services are provided to the client, 0 otherwise;
DELAY = number of days between financial year-end and the date of the audit report, a measure of the extent of audit report lag;
GEODISP = 1 if the client is geographically dispersed, 0 otherwise;
ORGCOMPL = 1 if the client's organizational structure as assessed by the audit team is complex, 0 otherwise;
SUBYN = 1 if the client has subsidiaries, 0 otherwise;
TIMING = 1 if extensive interim work is performed, 0 otherwise;
AUTOM = 1 if the client is extensively automated, 0 otherwise;
ICRELY = 1 if the audit team indicates high reliance on the client's internal controls, 0 otherwise;
FEERESID = residual from an audit fee model estimated using the companies in the sample;
PUBLIC = 1 if the client is publicly listed, 0 otherwise; and
YEAREND = 1 if the client has a December year-end, 0 otherwise.

TABLE 5

Tobit Regression of *LNEFF* on Client and Engagement Characteristics

Variable	Coeff. Est.	t-ratio	p-value
Intercept	−0.235	−5.22	0.000
CATREV	0.011	1.84	0.067
NEWCLNT	−0.001	−0.03	ns.
TAXYN	−0.051	−1.99	0.048
MASYN	0.027	1.08	ns.
DELAY	0.000	1.42	ns.
GEODISP	−0.021	−0.92	ns.
ORGCOMPL	0.024	0.95	ns.
SUBYN	−0.047	−1.90	0.058
TIMING	0.024	0.88	ns.
AUTOM	0.053	2.19	0.029
ICRELY	−0.059	−2.48	0.014
FEERESID	−0.008	−0.53	ns.
PUBLIC	0.020	0.70	ns.
YEAREND	0.040	1.77	0.078
OFF15	−0.097	−2.46	0.015
OFF247	−0.675	−6.10	0.000
OFF271	−0.241	−2.24	0.026
Log likelihood		58.167	
Likelihood ratio test statistic		71.93	0.000
Pseudo R^2		−1.620	
Number of observations used (n)		226	

Variable Definitions:

LNEFF = natural logarithm of efficiency score;
CATREV = size of the client measured by revenues, classified into 11 categories;
NEWCLNT = 1 if the duration of auditor tenure is three years or less, 0 otherwise;
TAXYN = 1 if tax services are provided to the client, 0 otherwise;
MASYN = 1 if consulting services are provided to the client, 0 otherwise;
DELAY = number of days between financial year-end and the date of the audit report, a measure of the extent of audit report lag;
GEODISP = 1 if the client is geographically dispersed, 0 otherwise;
ORGCOMPL = 1 if the client's organizational structure as assessed by the audit team is complex, 0 otherwise;
SUBYN = 1 if the client has subsidiaries, 0 otherwise;
TIMING = 1 if extensive interim work is performed, 0 otherwise;
AUTOM = 1 if the client is extensively automated, 0 otherwise;
ICRELY = 1 if the audit team indicates high reliance on the client's internal controls, 0 otherwise;
FEERESID = residual from an audit fee model estimated using the companies in the sample;
PUBLIC = 1 if the client is publicly listed, 0 otherwise;
YEAREND = 1 if the client has a December year-end, 0 otherwise;
OFF15 = 1 if the client is a client of audit firm office nr. 15, 0 otherwise;
OFF247 = 1 if the client is a client of audit firm office nr. 247, 0 otherwise; and
OFF271 = 1 if the client is a client of audit firm office nr. 271, 0 otherwise.

subsidiaries. The negative coefficient on *TAXYN* is interesting because it suggests that auditors may be less efficient when the firm is also providing tax services to a client, possibly because of increased risk or complexity.

The loss of efficiency attributable to relying on internal control is counter to professional guidance. However, this result is consistent with other research that has either found no relationship between internal control reliance and audit fees (O'Keefe, et al. 1994; Felix et al. 2001), or found a positive relationship between reliance on internal control and the level of auditor effort (Stein et al. 1994) or fees charged (Stein et al. 1994; Hay et al. 2006). In other words, while the audit risk model suggests that an auditor should trade off testing of internal controls for substantive tests when it is efficient, the empirical evidence indicates auditors are not effective in doing so. This may be because auditors are uncomfortable with "soft" evidence from internal controls and do not place enough weight on the significance of the results from tests of controls. As a result, auditors may follow a fairly standard approach to a specific audit regardless of the espoused reliance on internal control (Waller 1993). Further examination of our data reveals that the average cost per hour of engagements where the auditor relies on internal control is $99.24, while the average cost when auditors do not rely on internal control is $98.47. While the difference would appear to be slight, it hints that "reliance" audits may require more experienced personnel on the engagement team. In any event, given the prominent focus on internal control over financial reporting that is evident in the Public Company Accounting Oversight Board's (PCAOB) view of "the integrated audit," this result suggests that audits could become less efficient as a result of the auditing standards that heavily emphasize the role of internal control over financial reporting.[19] Alternatively, increased focus on internal control might force auditors to become more efficient at incorporating the results of risk assessment and test of controls in their evaluation of the overall evidence obtained during an audit.

Interestingly, *FEERESID* and *MASYN* are *not* significant in Table 5. Since research has suggested that efficiencies may be gained from knowledge spillovers when nonaudit services are jointly provided with an audit (Solomon 1990), the lack of statistical significance for *MASYN*, combined with our negative results for *TAXYN*, brings that view into question. The fact that *FEERESID* is not significant suggests that unusual levels of audit fees as measured by unexplained residuals from a fee-prediction model are not related to audit efficiency. In other words, on average, auditors do not charge more for inefficient audits, which is consistent with audit markets being competitive. Nor do auditors charge less for highly efficient audits, which suggest that clients are either unaware of the auditor's efficiency gains, or do not have sufficient negotiating power to obtain lower fees when an audit requires relatively less effort.

Potential Proxies for Audit Efficiency

Audit efficiency is an important topic for audit research because it may be associated with differential levels of audit quality. To assist other researchers, we now examine three possible proxies for audit efficiency: (1) residuals from a regression model for total audit hours, (2) audit report lag, and (3) residuals from a regression model for audit fees.

(1) Audit Hours Residuals and Regression-Based Production Models for Total Hours

Most previous research into audit production has used a regression-based approach for identifying factors that influence auditor effort. Table 6 reports the results of a traditional production model using two formulations: (1) the model of production originally reported in Hackenbrack and Knechel (1997), and (2) a model consistent with our second-stage analysis of efficiency scores. Each model uses the natural logarithm of total audit hours

---

[19] See Auditing Standard No. 2, *Auditing Internal Control Over Financial Reporting*, and Auditing Standard No. 5, *An Audit of Internal Control Over Financial Reporting that is Integrated with an Audit of Financial Statements*, for more details (PCAOB 2004, 2007). Public perception is that the cost of complying with Section 404 is quite high (Hartman 2005).

TABLE 6
OLS Regression of *LNTOTHRS* on Various Client and Engagement Characteristics Including Efficiency Scores

Panel A: Audit Hour Regression Based on Model in Hackenbrack and Knechel (1997)

Variable	Expected Sign	Coeff. Est.	t-ratio	p-value	Coeff. Est.	t-ratio	p-value
Intercept	?	4.742	34.49	0.000	4.866	31.91	0.000
CATREV	+	0.233	11.07	0.000	0.228	10.79	0.000
MASYN	−	0.246	2.73	0.007	0.231	2.56	0.011
SUBYN	+	0.398	4.41	0.000	0.405	4.51	0.000
ICRELY	−	0.086	0.96	ns.	0.106	1.19	ns.
PUBLIC	+	0.402	3.99	0.000	0.399	3.98	0.000
FININD	−	−0.329	−3.93	0.000	−0.351	−4.17	0.000
LNEFF	−				0.483	1.84	0.067
Overall F-test			74.07	0.000		64.66	0.000
Adjusted R^2			0.661			0.665	
Number of observations used (n)			226			226	

Panel B: Audit Hour Regression Based on Second-Stage Model in Table 5

Variable	Expected Sign	Coeff. Est.	t-ratio[a]	p-value	Coeff. Est.	t-ratio[a]	p-value
Intercept	?	4.725	45.38	0.000	4.758	41.18	0.000
CATREV	+	0.228	13.57	0.000	0.228	13.21	0.000
NEWCLNT	−	0.048	0.69	ns.	0.047	0.67	ns.
TAXYN	−	−0.131	−2.15	0.033	−0.127	−2.09	0.038
MASYN	−	0.232	3.22	0.001	0.229	3.14	0.002
DELAY	+	0.001	1.81	0.072	0.001	1.75	0.081
GEODISP	+	−0.017	−0.36	ns.	−0.015	−0.32	ns.
ORGCOMPL	+	0.124	2.03	0.044	0.122	1.99	0.048
SUBYN	+	0.369	5.86	0.000	0.371	5.83	0.000
TIMING	−	0.094	1.03	ns.	0.093	1.02	ns.
AUTOM	−	−0.026	−0.36	ns.	−0.032	−0.45	ns.
ICRELY	−	0.045	0.65	ns.	0.052	0.74	ns.
FEERESID	+	0.628	5.75	0.000	0.627	5.73	0.000
PUBLIC	+	0.372	5.36	0.000	0.369	5.34	0.000
YEAREND	+	−0.126	−2.53	0.012	−0.132	−2.63	0.009
LNEFF	−				0.127	0.82	ns.
Overall F-test			156.27	0.000		147.88	0.000
Adjusted R^2			0.854			0.853	
Number of observations used (n)			226			226	

---

[a] t-statistics are calculated using White's correction for heteroscedasticity.

Variable Definitions:
LNTOTHRS = natural logarithm of total audit hours;
CATREV = size of the client measured by revenues, classified into 11 categories;
NEWCLNT = 1 if the duration of auditor tenure is three years or less, 0 otherwise;
TAXYN = 1 if tax services are provided to the client, 0 otherwise;
MASYN = 1 if consulting services are provided to the client, 0 otherwise;

(*continued on next page*)

TABLE 6 (continued)

DELAY = number of days between financial year-end and the date of the audit report, a measure of the extent of audit report lag;
GEODISP = 1 if the client is geographically dispersed, 0 otherwise;
ORGCOMPL = 1 if the client's organizational structure as assessed by the audit team is complex, 0 otherwise;
SUBYN = 1 if the client has subsidiaries, 0 otherwise;
TIMING = 1 if extensive interim work is performed, 0 otherwise;
AUTOM = 1 if the client is extensively automated, 0 otherwise;
ICRELY = 1 if the audit team indicates high reliance on the client's internal controls, 0 otherwise;
FEERESID = residual from an audit fee model estimated using the companies in the sample;
PUBLIC = 1 if the client is publicly listed, 0 otherwise;
YEAREND = 1 if the client has a December year-end, 0 otherwise;
FININD = 1 if the client operates in the finance, insurance or real estate industry, 0 otherwise; and
LNEFF = natural logarithm of efficiency score.

(*LNTOTHRS*) as the dependent variable and is estimated with and without an independent variable for efficiency. We can examine the relationship between efficiency and traditional measures of auditor effort since our measures of efficiency are derived independently of the client characteristics. We use the natural log of our efficiency scores (*LNEFF*) as a proxy for production efficiency and estimate the following models:

$$\begin{aligned} LNTOTHRS = {} & \beta_0 + \beta_1 CATREV + \beta_2 MASYN + \beta_3 SUBYN + \beta_4 ICRELY \\ & + \beta_5 PUBLIC + \beta_6 FININD + \beta_7 LNEFF + \varepsilon; \end{aligned} \tag{2}$$

$$\begin{aligned} LNTOTHRS = {} & \beta_0 + \beta_1 CATREV + \beta_2 NEWCLNT + \beta_3 TAXYN + \beta_4 MASYN \\ & + \beta_5 DELAY + \beta_6 GEODISP + \beta_7 ORGCOMPL + \beta_8 SUBYN \\ & + \beta_9 TIMING + \beta_{10} AUTOM + \beta_{11} ICRELY + \beta_{12} FEERESID \\ & + \beta_{13} PUBLIC + \beta_{14} YEAREND + \beta_{15} LNEFF + \varepsilon. \end{aligned} \tag{3}$$

Results are reported in Table 6. The Hackenbrack and Knechel model (Equation (2), Panel A) has an adjusted R^2 of 0.661 ($p < 0.000$) without the efficiency proxy. When *LNEFF* is added to the model, it is significant and positive ($p < 0.067$) and the adjusted R^2 increases to 0.665 ($p < 0.000$). This result implies that audits that consume more total hours are generally more efficient. While this result may seem surprising, it is consistent with the observation from prior research that the use of lower-cost staff time increases disproportionately as total hours increase. In Panel B (Equation (3)), our more extensive model of production has an adjusted R^2 of 0.854 ($p < 0.000$) without the efficiency proxy. When *LNEFF* is added to the model, it is positive but not significant, and the adjusted R^2 remains virtually the same.[20] These results suggest that a fully developed regression model of audit production captures the effect of many of the sources of inefficiency in an audit.

(2) Audit Report Lag

Knechel and Payne (2001) observe that audit report lag is associated with the unexplained residuals of a regression-based model of total audit hours as long as certain client characteristics are incorporated into the regression analysis. They suggest that audit report lag might be interpreted as a measure of audit efficiency that can be externally observed for any company. As reported in Table 5, we find no significant relationship between audit report lag and audit efficiency when lag is treated as an independent variable (*DELAY*). For the subsample of audits that are inefficient, audit report lag is positively associated with efficiency ($p < 0.046$). These results suggest that audit report lag may be a weak indicator of audit efficiency.

---

[20] In line with these findings, further examination of the residuals from the production models without the efficiency proxies indicates that the correlation between efficiency and the residuals from the Hackenbrack and Knechel model is 0.120 ($p < 0.073$) while the correlation between efficiency and the residuals from the current study's production model is 0.050 ($p < 0.452$).

We further examine the link between audit report lag and efficiency by replicating the analysis in Knechel and Payne (2001) and adding our efficiency score to their audit report lag model. Since their model incorporates the residual from an audit hour regression model, we conduct this analysis using both models reported in Table 6. Our results (not reported) indicate that the coefficient on audit efficiency is positive but not significant in either case. These results suggest that while audit report lag can be a proxy for unusual audit effort (per Hackenbrack and Knechel's [1997] findings), it may be a weak external proxy for audit efficiency per se.

(3) Audit Fee Residuals

Another potential proxy for audit efficiency is the residual from an audit fee model. In a competitive market, an auditor should not be able to pass on the cost of inefficiencies in the production of an audit since other competitors would be able to offer clients a lower fee by being more efficient. That is, while inefficiency may occur in any process, it should not, on average, be associated with higher fees. As previously noted, the audit fee residual is not significant in Table 5 (*FEERESID*). Also, when our efficiency score is included in a model to predict audit fees the efficiency score is not significant (not reported). Thus, we find no evidence that the residual for an audit fee model is a proxy for audit efficiency in a pooled analysis. To provide more insight into this issue, we consider how *FEERESID* interacts with either *INDUSTRY* or *OFFICE* to determine if clusters of audit services reveal an association between efficiency and fees. To this end, we estimate the following models:

$$LNEFF = \beta_0 + \beta_1 CATREV + \beta_{2-8}(FEERESID * INDUSTRY) + \varepsilon; \tag{4}$$

$$LNEFF = \beta_0 + \beta_1 CATREV + \beta_2 FEERESID + \beta_{3-17}(FEERESID * OFFICE) + \varepsilon. \tag{5}$$

In Panel A of Table 7, we omit the main effect for *FEERESID* and analyze each industry separately. We see that the association between efficiency and fee residuals is significantly positive for clients in the manufacturing industry (*FEERESID* * *MANIND*, $p < 0.024$) and significantly negative for clients in the health care industry (*FEERESID* * *HCIND*, $p < 0.047$). The former suggests that fees are higher when audits are efficient in the manufacturing industry, while the latter indicates that fees are higher when the auditor has inefficient audits in the health care industry. Since manufacturing is generally considered to be a less specialized area for auditors, the positive association between efficiency and fee residuals suggests that auditors may have an incentive to be highly efficient in such audits. However, the negative association with fee residuals in health care audits, a specialized area of auditing (e.g., Mayhew and Wilkins 2003), may indicate that the greater the need for a specialist auditor, the less competitive the audit market. That is, the ability to price audit inefficiency in some industries may imply that specialized areas of auditing may be less competitive.

TABLE 7
Analysis of Fees Conditional on Industry and Office Location

Panel A: Tobit Regression of *LNEFF* on *CATREV* and *FEERESID * INDUSTRY*

Variable	Coeff. Est.	t-ratio	p-value
Intercept	−0.242	−6.27	0.000
CATREV	0.010	2.10	0.037
*FEERESID * MANIND*	0.110	2.27	0.024
*FEERESID * FININD*	0.021	0.79	ns.
*FEERESID * GOVIND*	0.108	1.40	ns.
*FEERESID * HCIND*	−0.107	−2.00	0.047
*FEERESID * HTIND*	0.078	1.08	ns.
*FEERESID * MERIND*	0.025	0.53	ns.
*FEERESID * EDUIND*	−0.043	−1.02	ns.
Log likelihood		32.161	
Likelihood ratio test statistic		19.92	0.011
Pseudo R^2		−0.449	
Number of observations used (*n*)		226	

Panel B: Tobit Regression of *LNEFF* on *CATREV, FEERESID,* and *FEERESID * OFFICE*

Variable	Coeff. Est.	t-ratio	p-value
Intercept	−0.258	−6.52	0.000
CATREV	0.012	2.50	0.013
FEERESID	0.073	1.89	0.060
*FEERESID * OFF15*	0.032	0.33	ns.
*FEERESID * OFF69*	−0.072	−0.75	ns.
*FEERESID * OFF75*	−0.057	−0.57	ns.
*FEERESID * OFF105*	−0.145	−1.51	ns.
*FEERESID * OFF115*	−0.143	−1.79	0.075
*FEERESID * OFF147*	−0.035	−0.46	ns.
*FEERESID * OFF163*	−0.080	−0.77	ns.
*FEERESID * OFF166*	0.075	0.80	ns.
*FEERESID * OFF181*	−0.076	−0.55	ns.
*FEERESID * OFF182*	−0.174	−1.78	0.076
*FEERESID * OFF194*	−0.117	−1.37	ns.
*FEERESID * OFF196*	−0.097	−1.77	0.078
*FEERESID * OFF204*	−0.044	−0.58	ns.
*FEERESID * OFF221*	−0.094	−0.83	ns.
*FEERESID * OFF251*	−0.057	−0.78	ns.
*FEERESID * OFF293*	0.046	0.50	ns.
Log likelihood		32.349	
Likelihood ratio test statistic		20.29	0.317
Pseudo R^2		−0.457	
Number of observations used (n)		226	

(*continued on next page*)

TABLE 7 (continued)

Variable Definitions:
LNEFF = natural logarithm of total audit hours;
CATREV = size of the client measured by revenues, classified into 11 categories;
FEERESID = residual from an audit fee model estimated using the companies in the sample;
MANIND = 1 if the client operates in the manufacturing industry, 0 otherwise;
FININD = 1 if the client operates in the finance, insurance, or real estate industry, 0 otherwise;
GOVIND = 1 if the client is a governmental organization, 0 otherwise;
HCIND = 1 if the client operates in the health care industry, 0 otherwise;
HTIND = 1 if the client operates in the high tech industry, 0 otherwise;
MERIND = 1 if the client operates in the merchandising industry, 0 otherwise;
EDUIND = 1 if the client operates in the educational sector, 0 otherwise;
OFF15 = 1 if the client is a client of audit firm office nr. 15, 0 otherwise;
OFF69 = 1 if the client is a client of audit firm office nr. 69, 0 otherwise;
OFF75 = 1 if the client is a client of audit firm office nr. 75, 0 otherwise;
OFF105 = 1 if the client is a client of audit firm office nr. 105, 0 otherwise;
OFF115 = 1 if the client is a client of audit firm office nr. 115, 0 otherwise;
OFF147 = 1 if the client is a client of audit firm office nr. 147, 0 otherwise;
OFF163 = 1 if the client is a client of audit firm office nr. 163, 0 otherwise;
OFF166 = 1 if the client is a client of audit firm office nr. 166, 0 otherwise;
OFF181 = 1 if the client is a client of audit firm office nr. 181, 0 otherwise;
OFF182 = 1 if the client is a client of audit firm office nr. 182, 0 otherwise;
OFF194 = 1 if the client is a client of audit firm office nr. 194, 0 otherwise;
OFF196 = 1 if the client is a client of audit firm office nr. 196, 0 otherwise;
OFF204 = 1 if the client is a client of audit firm office nr. 204, 0 otherwise;
OFF221 = 1 if the client is a client of audit firm office nr. 221, 0 otherwise;
OFF251 = 1 if the client is a client of audit firm office nr. 251, 0 otherwise; and
OFF293 = 1 if the client is a client of audit firm office nr. 293, 0 otherwise.

In Panel B of Table 7, we include the main effect for *FEERESID* but only include interaction terms for offices with at least ten engagements (i.e., the effect of all other offices is included in the model intercept). We see that three offices exhibit a significant negative association between efficiency and unexplained fees (office 115, $p < 0.075$; office 182, $p < 0.076$; office 196, $p < 0.078$). Since we do not know the actual locations of these offices we cannot determine if the association is due to specific attributes of the office (size, industry specialization, local competition), but the results do indicate that the ability to price variances in efficiency differs across office locations. This may be due to auditor concentration in those markets. Combined, the results in both panels suggest that analysis of audits based on industry specializations within offices may be worthwhile (Ferguson et al. 2003; Francis et al. 2006). Furthermore, the results suggest that audit fee studies that examine premiums for Big N firms or industry specialists may be detecting the effect of differences in the pricing of audit efficiency and competition.

V. CONCLUSION

As noted by Professor Kinney in remarks to the U.S. Treasury, a fundamental lack of understanding of auditing as a production process has caused "today's audit process [to be] treated as a 'black box' " (Kinney 2008). This has potentially stifled the improvement of audit practice. Our aim is to develop a theoretically valid, empirical model of audit production that considers the cost of labor hours as an input used to produce outputs in the form of evidence-gathering activities that increase the level of assurance obtained in an audit. Client-specific characteristics are considered exogenous factors in our model and are assumed to affect audit production as a whole.

We make three important theoretical contributions to research about audit production. First, we address how to translate the theoretical model of O'Keefe et al. (1994) that links effort (inputs) to achieved assurance (output) into a viable empirical model. We do this by defining auditor inputs as the cost of effort and auditor output as time spent on evidence-gathering activities that proxy for the achieved (unobservable) level of assurance. Second, this approach allows us to remove a critical assumption of prior research, namely, that all audits conducted by a firm within a given time period achieve a constant level of assurance. We do so by defining an output proxy that can vary with effort but is monotonically linked to the unobservable level of assurance. Third, we develop a technique to evaluate the trade-offs among evidence-gathering activities conducted by an auditor.

We apply our model empirically using a sample of U.S.-based audit engagements of an international accounting firm. We assess audit performance by computing efficiency scores using Data Envelopment Analysis. Our results suggest that a model based on variable returns to scale that includes some constraints on the level of certain evidence-gathering activities is the most appropriate. Furthermore, many audits in our sample are inefficient and efficiency varies across client industries and firm offices. The relatively low level of efficiency is partly due to the consideration of some audit procedures, such as non-critical testing, as being of low value. Because they provide less assurance than other activities, we incorporated constraints for these procedures in our empirical DEA model.

Next, we determine the effect of client-specific characteristics on audit efficiency. Our results show that the audits of large, automated clients with a year-end in December are generally more efficient. However, efficiency is reduced when a client has subsidiaries, the auditor relies on internal control, or provides tax services to a client. Interestingly, auditor tenure, provision of consulting services, organizational complexity, and listing status, which all have been shown to affect the *level* of auditor effort on an engagement, do not influence overall auditor efficiency. This suggests that the auditor properly calibrates many client attributes in determining the appropriate effort level for an engagement.

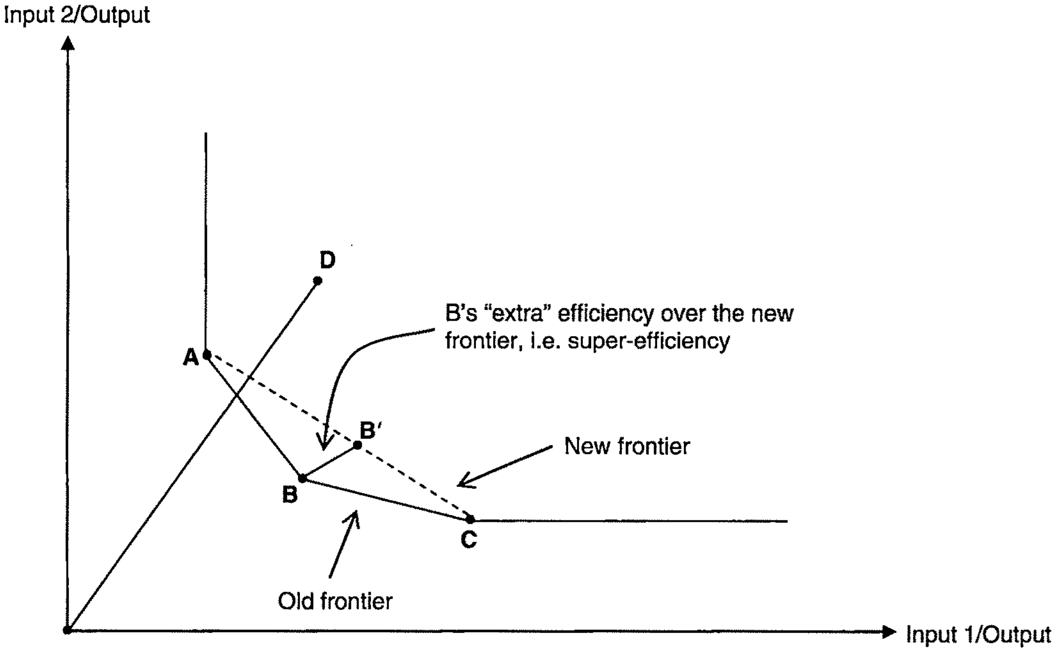
Some of our results may be of interest to empirical researchers in other areas. First, we find that conventional well-specified regression-based models of audit effort control for many of the factors that influence audit efficiency. However, we find little evidence that audit report lag or residuals from an audit fee regression model are associated with audit efficiency (unless calculated at an industry or office level). Thus, our results suggest that it may be quite difficult to assess audit efficiency without access to internal data for individual audits. Future research may need to have such access to further advance our knowledge of audit production (see Kinney 2008). Second, we obtain some interesting insights into the nature of audit markets. Classic models of audit fees assume that audit markets are competitive (Simunic 1980; Hay et al. 2006), which suggests that most audits must be conducted to a high level of efficiency since inefficient firms would be unlikely to survive. Our results suggest that this assumption of competitive markets may be difficult to defend at the industry/location level of analysis.

Some limitations should be taken into account when interpreting our results. First, the data come from a single accounting firm and a single time period in the early 1990s. Clearly, the age of the data may affect the extent to which the results can be generalized in today's market for audit services. Data from more recent audits, especially after the passage of the Sarbanes-Oxley Act of 2002, might reveal different causes of audit efficiency, particularly with respect to the testing of internal controls. A recent study by Bell et al. (2008) indicates that total labor used in audits conducted in 2002 is marginally lower than audits in the 1990s, but that the amount of higher-ranked labor increased significantly. Second, the presumed relationship between the evidence-gathering activities and the final unobservable outcome (achieved assurance) may be subject to measurement error. Third, audits also consume "overhead" as part of the audit process but we have no data on the capital investment needed to conduct large international audits. Fourth, some of the variables used in the regression model are based on post-audit assessments by the audit team, which introduces potential measurement bias in these variables. The model and approach that we use offer considerable potential for modeling audit production as well as understanding the factors that affect audit efficiency. The clear separation between the direct audit activities and resources *vis-à-vis* external factors is intuitive and corresponds closely to the production economics and management accounting literature. We believe that these advantages and our approach should be considered by those continuing to expand research in this area.

APPENDIX A
AN ILLUSTRATION OF DEA

Figure 3 illustrates the use of DEA by depicting four DMUs, A, B, C and D, that produce a single output and consume two inputs. Minimum input combinations lie on the frontier connecting A, B, and C since no other DMU produces the same output with a lower input combination. In contrast, unit D is dominated by the other DMUs because it produces the same output with a higher input combination. D's inefficiency can be measured by the radial distance to the frontier along the ray extending from the origin to D and intersecting the AB segment of the frontier. DMUs on the efficiency frontier are identified as fully efficient and assigned an efficiency score of 100 percent. Inefficient DMUs are assigned scores of less than 100 percent.

FIGURE 3
An Example of DEA with Two Inputs and One Output for Four DMUs A, B, C, and D


Input 2/Output
D
B's "extra" efficiency over the new frontier, i.e. super-efficiency
A
B'
New frontier
B
C
Old frontier
Input 1/Output

Further analysis of the efficient set of DMUs is possible by computing efficiency scores in excess of 100 percent. These are so-called super-efficiency scores. If unit B were to be excluded from the frontier, then a new frontier would be created comprising DMUs A and C. B's super-efficient score is the distance to the new frontier where this "extra" efficiency represents the allowable increase in its inputs before it would become inefficient. For example, a score of 120 percent for DMU B would imply that it could increase its inputs by 20 percent (to B′) and still remain efficient. This so-called super-efficient model (Andersen and Petersen 1993; Banker et al. 1989) is applied in our analyses using the approach described in Lovell and Rouse (2003). Following Banker and Chang (2006), we use the super-efficiency scores to identify possible outliers, resulting in two audits with super-efficiency scores in excess of 150 percent being dropped from our analysis.

To further illustrate how DEA calculates efficiency for a given unit (audit), consider the following simplified example:[21]

	Audit A	**Audit B**
Inputs		
Partner cost	$1,000	$500
Staff cost	$2,000	$1,500
Output		
Substantive testing hours	10	10

Audit A is less efficient than Audit B, having cost more in partner and staff time. In fact, B has performed the same level of substantive testing for 50 percent less partner cost and 25 percent less staff cost. Based on staff cost, Audit A is 25 percent less efficient than Audit B while it is 50 percent less efficient in terms of partner cost. DEA seeks to maximize a unit's efficiency so the model would assign an efficiency score of 75 percent to Audit A (i.e., 100 percent minus the smaller of the two inefficiencies). Audit B's efficiency would be scored at 100 percent.

The reduction of Audit A's inputs by 25 percent would equate the staff cost to that of Audit B, but would still leave a difference of $250 ($1,000 * 75 percent − $500) for partner cost. This residual is known as non-radial slack and represents the difference between strong and weak efficiency. DEA efficiency scores therefore measure equi-proportionate reductions in inputs (or increases in outputs) but do not always ensure strong efficiency. This equi-proportionate movement is called a radial distance measure of (in)efficiency. This can be seen more clearly in an extension of the above example that includes audits C and D.

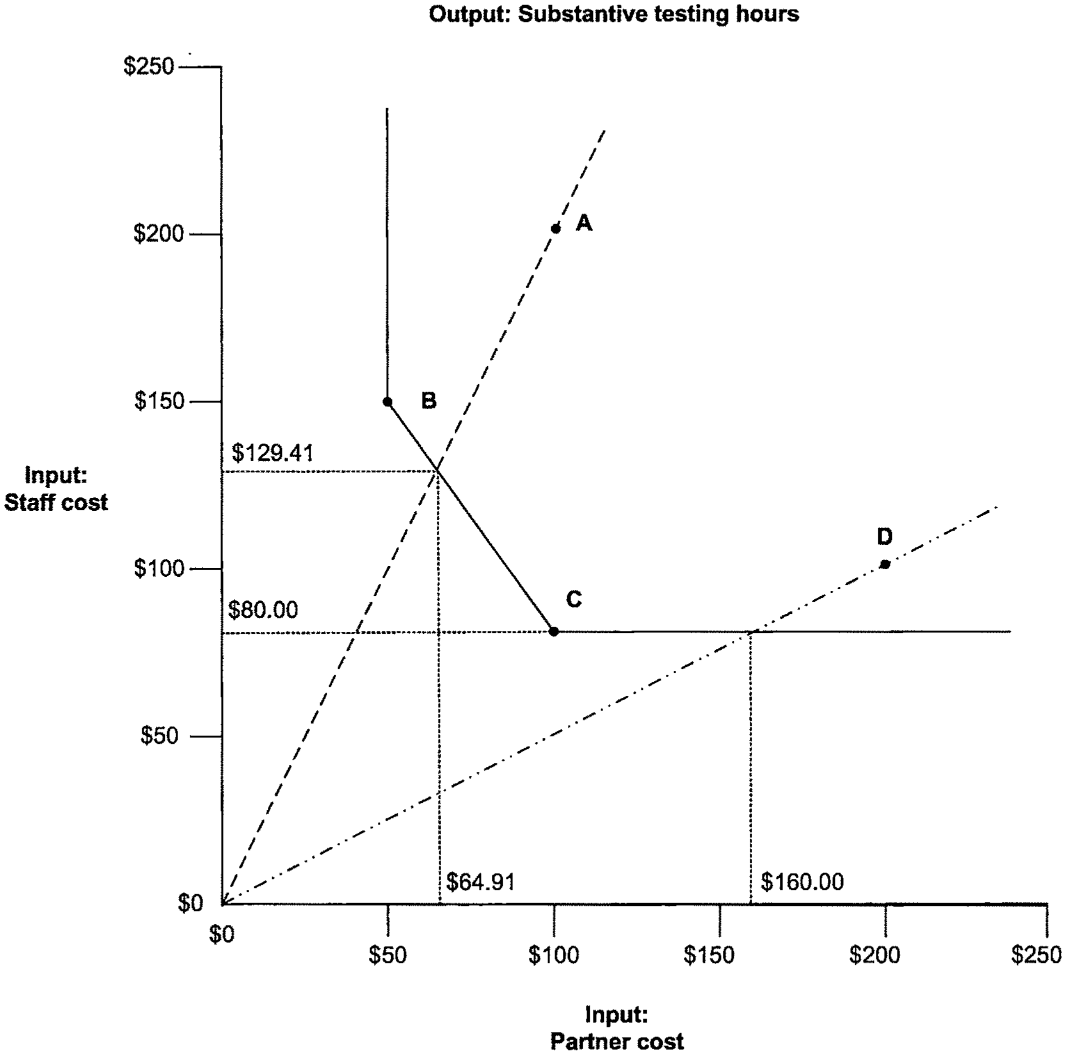
	Audit A	**Audit B**	**Audit C**	**Audit D**
Inputs				
Partner cost	$1,000	$500	$1,000	$2,000
Staff cost	$2,000	$1,500	$800	$1,000
Outputs				
Substantive testing hours	10	10	10	10

Each audit is plotted in Figure 4 in terms of the input cost per hour. Audits B and C form an efficient frontier. Audits A and D have higher cost combinations and are dominated

---

[21] Note: The data are drawn from an example provided by Mike Stein in his discussant's presentation at the 2005 Auditing Section Midyear Conference.

FIGURE 4
An Extended Example of DEA with Two Inputs and One Output for Four DMUs A, B, C, and D



by this frontier. The level of inputs required to make Audit A efficient are indicated by the intersection of the frontier connecting B and C with the ray drawn from the origin to A (0A). This point of intersection is $64.91 partner cost and $129.41 staff cost. Dividing these by A's actual input levels provides its radial efficiency score of 64.7 percent. Therefore, if A reduces its partner and staff costs equi-proportionately (i.e., radially) by 35.3 percent, then it will reach the frontier at the point of intersection indicated by the 0A ray.

Audit D is slightly different. Although its inputs can be reduced radially to the point of intersection (160, 80) with the frontier that extends horizontally from C, this point is still dominated by audit C (100, 80). The horizontal shift sideways from $160 partner cost to \$100 partner cost is non-radial slack. Therefore D's radial measure of efficiency (80 percent) will move it to the weakly efficient part of the frontier but a non-radial reduction is still required to project it to the strong efficient frontier. While the distinction between strong and weak efficiency is conceptually important, it is not usually a problem in practical applications. Nonetheless, non-radial slack provides a good indication of situations where the frontier does not envelop the data very well or where there are instances of weakly efficient units.

APPENDIX B
FORMAL SPECIFICATION OF THE DEA MODEL USED IN THE PRIMARY ANALYSIS

The general form of the input-oriented CRS DEA envelopment and multiplier models are specified as follows:

Envelopment Model	**Multiplier Model**
Min θ_o	Max $\boldsymbol{\mu}^T \boldsymbol{y}_o$
s.t. $\boldsymbol{Y\lambda} \geq \boldsymbol{y}_o$	s.t. $\boldsymbol{\mu}^T \boldsymbol{Y} - \boldsymbol{v}^T \boldsymbol{X} \leq 0$
$\boldsymbol{X\lambda} \leq \boldsymbol{x}_o \theta_o$	$\boldsymbol{v}^T \boldsymbol{x}_o = 1$
$\boldsymbol{\lambda}, \geq 0$; θ_o free	$\boldsymbol{\mu}, \boldsymbol{v} \geq 0$

The notation is for outputs $(y_1, y_2, \ldots, y_s)$ and inputs $(x_1, x_2, \ldots, x_m)$ for DMUs $j = 1, 2, \ldots, n$. $\boldsymbol{Y}$ is an $s \times (n)$ matrix of outputs, $\boldsymbol{X}$ is an $m \times (n)$ matrix of inputs, and $\boldsymbol{\lambda}$ is an (n)-dimensional vector of intensity variables for DMUs j. $\boldsymbol{y}_o$ and $\boldsymbol{x}_o$ are output and input vectors for DMU$_o$ being evaluated. In this audit application, there are 307 DMUs (n = 307), eight outputs, and one input. The formulation for the multiplier model is set out below with weight constraints included. In total there are 311 constraints (307 + 1 + 3) excluding the implicit non-negative constraints.

$$\text{Max } \mu_1 y_{1o} + \mu_2 y_{2o} + \mu_3 y_{3o} + \mu_4 y_{4o} + \mu_5 y_{5o} + \mu_6 y_{6o} + \mu_7 y_{7o} + \mu_8 y_{8o}$$

subject to:

$$\mu_1 y_{1j} + \mu_2 y_{2j} + \mu_3 y_{3j} + \mu_4 y_{4j} + \mu_5 y_{5j} + \mu_6 y_{6j} + \mu_7 y_{7j} + \mu_8 y_{8j} - \upsilon_1 x_{1j} \leq 0$$

where $j = 1, \ldots, 307$

$$\begin{aligned} \upsilon_1 x_{1o} &= 1 \\ \mu_3 - s\,\mu_4 &\geq 0 \\ \mu_5 - s\,\mu_6 &\geq 0 \\ \mu_3 - \mu_5 &\geq 0 \\ \mu, \upsilon &\geq 0. \end{aligned}$$

In this model, subscript "o" denotes the audit under evaluation. The following notation corresponds to our data:

y_1 denotes output (1) audit planning and $\boldsymbol{\mu}_1$ is the associated weight;
y_2 denotes output (2) internal control evaluation and $\boldsymbol{\mu}_2$ is the associated weight;

y_3 denotes output (3) Substantive Testing—Critical objectives and μ_3 is the associated weight;

y_4 denotes output (4) Substantive Testing—Non-Critical objectives and μ_4 is the associated weight;

y_5 denotes output (5) Review—Critical and μ_5 is the associated weight;

y_6 denotes output (6) Review—Non-Critical and μ_6 is the associated weight;

y_7 denotes output (7) financial statement preparation and μ_7 is the associated weight;

y_8 denotes output (8) client interaction and μ_8 is the associated weight;

x_1 denotes the single input comprising total labor costs and υ_1 is the associated weight; and

s denotes a scalar to reflect the relative trade-off in value of the paired activities.

REFERENCES

Andersen, P., and N. C. A. Petersen. 1993. A procedure for ranking efficiency units in Data Envelopment Analysis. *Management Science* 39 (10): 1261–1264.

Banker, R. D., S. Das, and S. M. Datar 1989. Analysis of cost variances for management control in hospitals. *Research in Governmental and Non-profit Accounting* 5: 268–291.

———. 1993. Maximum likelihood, consistency and Data Envelopment Analysis: A statistical foundation. *Management Science* 39 (10): 1265–1273.

———, and R. Natarajan. 2004. Statistical tests based on DEA efficiency scores. In *Handbook on Data Envelopment Analysis*, edited by Cooper, W. W., L. Seiford, and J. Zhu, Chap. 11, 299–321. Norwell, MA: Kluwer Academic Publishers.

———, and H. Chang. 2006. The super-efficiency procedure for outlier identification, not for ranking efficient units. *European Journal of Operational Research* 175 (2/1): 1311–1320.

Bell, T., R. Doogar, and I. Solomon. 2008. Audit labor usage and fees under business risk auditing. *Journal of Accounting Research* 46 (4): 729–760.

Cooper, W. W., L. M. Seiford, and K. Tone. 2007. *Data Envelopment Analysis: A Comprehensive Text with Models, Applications, References and DEA-Solver Software.* Second edition. New York, NY: Springer Science + Business Media, LLC.

Davidson, R. A., and W. E. Gist. 1996. Empirical evidence on the functional relation between audit planning and total audit effort. *Journal of Accounting Research* 34 (1): 111–124.

Davis, L., D. Ricchiute, and G. Trompeter. 1993. Audit effort, audit fees, and the provision of nonaudit services to audit clients. *The Accounting Review* 68 (1): 135–150.

Dopuch, N., M. Gupta, D. Simunic, and M. Stein. 2003. Production efficiency and the pricing of audit services. *Contemporary Accounting Research* 20 (1): 79–115.

Färe, R., and S. Grosskopf. 2004. Modeling undesirable factors in efficiency evaluation: Comment. *European Journal of Operational Research* 157 (1): 242–245.

Felix, W. L., Jr., A. A. Gramling, and M. J. Maletta. 2001. The contribution of internal audit as a determinant of external audit fees and factors influencing this contribution. *Journal of Accounting Research* 39 (3): 513–534.

Ferguson, A., J. R. Francis, and D. J. Stokes. 2003. The effects of firm-wide and office-level industry expertise on audit pricing. *The Accounting Review* 78 (2): 429–448.

Francis, J. R., K. Reichelt, and D. Wang. 2006. National versus office-specific measures of auditor industry expertise and effects on client earnings quality. Working paper, University of Missouri.

———, and M. D. Yu. 2009. Big 4 office size and audit quality. *The Accounting Review* 84 (5): 1521–1552.

Fried, H. O., C. A. K. Lovell, S. S. Schmidt, and S. Yaisawarng. 2002. Accounting for environmental effects and statistical noise in Data Envelopment Analysis. *Journal of Productivity Analysis* 17 (1–2): 157–174.

Gujarati, D. N. 1995. *Basic Econometrics*. 3rd edition. New York, NY: McGraw-Hill, Inc.

Hackenbrack, K., and W. R. Knechel. 1997. Resource allocation decisions in audit engagements. *Contemporary Accounting Research* 14 (3): 481–499.

Hartman, T. E. 2005. *The Cost of Being Public in the Era of Sarbanes-Oxley*. West Palm Beach, FL: Foley and Lardner LLP.

Hay, D., W. R. Knechel, and N. Wong 2006. Audit fees: A meta-analysis of the effect of supply and demand attributes. *Contemporary Accounting Research* 23 (1): 141–191.

———, and ———. 2009. The effects of advertising and solicitation on audit fees. *Journal of Accounting and Public Policy* (forthcoming).

Jensen, K., J. M. Kim, and H. Yi. 2008. The geography of U.S. auditors. Working paper, The University of Oklahoma.

Kinney, W. R. 2008. The human capital role of independent scholarship in auditing. Prepared remarks for testimony before the U. S. Department of the Treasury. Washington, D.C.: Government Printing Office.

Knechel, W. R., and J. L. Payne. 2001. Additional evidence on audit report lag. *Auditing: A Journal of Practice & Theory* 20 (1): 137–146.

Lovell, C. A. K., and A. P. B. Rouse. 2003. Equivalent standard DEA models to provide super-efficiency scores. *The Journal of the Operational Research Society* 54 (1): 101–108.

Mayhew, B. W., and M. S. Wilkins. 2003. Audit firm industry specialization as a differentiation strategy: Evidence from fees charged to firms going public. *Auditing: A Journal of Practice & Theory* 22 (2): 33–52.

O'Keefe, T. B., D. A. Simunic, and M. T. Stein. 1994. The production of audit services: Evidence from a major public accounting firm. *Journal of Accounting Research* 32 (2): 241–261.

Palmrose, Z.-V. 1986. Audit fees and auditor size: further evidence. *Journal of Accounting Research* 24 (1): 97–110.

———. 1989. The relation of audit contract type to audit fees and audit hours. *The Accounting Review* 64 (3): 488–499.

Pitt, M. M., and L.-F. Lee. 1981. Measurement and sources of technical inefficiency in the Indonesian weaving industry. *Journal of Development Economics* 9 (1): 43–64.

Public Company Accounting Oversight Board (PCAOB). 2004. *Auditing Internal Control Over Financial Reporting Performed in Conjunction with an Audit of Financial Statements*. Auditing Standard No. 2. Washington, D.C.: PCAOB.

———. 2007. *An Audit of Internal Controls Over Financial Reporting that is Integrated with an Audit of Financial Statements*. Auditing Standard No. 5. Washington, D.C.: PCAOB.

Reifschneider, D., and R. Stevenson. 1991. Systematic departures from the frontier: Framework for the analysis of firm inefficiency. *International Economic Review* 32 (3): 715–723.

Rouse, P., M. Putterill, and D. Ryan. 1997. Towards a general managerial framework for performance measurement: A comprehensive highway maintenance application. *Journal of Productivity Analysis* 8 (2): 127–149.

Seiford, L. M., and J. Zhu. 2002. Modeling undesirable factors in efficiency evaluation. *European Journal of Operational Research* 142 (1): 16–20.

Shepherd, R. W. 1970. *Theory of Cost and Production Functions*. Princeton, NJ: Princeton University Press.

Simunic, D. A. 1980. The pricing of audit services: Theory and evidence. *Journal of Accounting Research* 18 (1): 161–190.

Solomon, I. 1990. Discussion of: The jointness of audit fees and demand for MAS: a self-selection analysis. *Contemporary Accounting Research* 6 (2): 323–328.

Stein, M. T., D. A. Simunic, and T. B. O'Keefe. 1994. Industry differences in the production of audit services. *Auditing: A Journal of Practice & Theory* 13 (Supplement): 128–142.

Waller, W. 1993. Auditors' assessments of inherent and control risk in field settings. *The Accounting Review* 68 (4): 783–803.

THE ACCOUNTING REVIEW
Vol. 84, No. 5
2009
pp. 1639–1670

American Accounting Association
DOI: 10.2308/accr.2009.84.5.1639

The Effect of Disclosures by Management, Analysts, and Business Press on Cost of Capital, Return Volatility, and Analyst Forecasts: A Study Using Content Analysis

S. P. Kothari
Massachusetts Institute of Technology and Barclays Global Investors, San Francisco

Xu Li
The University of Texas at Dallas

James E. Short
University of California, San Diego

ABSTRACT: We document systematic evidence of risk effects of disclosures culled from a virtually exhaustive set of sources from the print medium. We content analyze more than 100,000 disclosure reports by management, analysts, and news reporters (i.e., business press) in constructing firm-specific disclosure measures that are quantitative and amenable to replication. We expect credibility and timeliness differences in the disclosures by source, which would translate into differential cost of capital effects. We find that when content analysis indicates favorable disclosures, the firm's risk, as proxied by the cost of capital, stock return volatility, and analyst forecast dispersion, declines significantly. In contrast, unfavorable disclosures are accompanied by significant increases in risk measures. Analysis of disclosures by source—corporations, analysts, and the business press—reveals that negative disclosures from business press sources result in increased cost of capital and return volatility, and favorable reports from business press reduce the cost of capital and return volatility.

Keywords: *disclosure; cost of capital; content analysis.*

We are grateful to Kalani Oshiro, Stanley Woods, Mitul Mehta, and Tanis Fidelholtz for their research assistance. We have benefited from discussions and comments from Philip Wright, Wayne Guay, Joel Kurtzman, Mark Lutchen, Sunil Misser, Alison Thomas, and Gabrielle Wong. We acknowledge financial support from PricewaterhouseCoopers and the MIT Center for eBusiness.

Editor's note: Accepted by Steven Kachelmeier, with thanks to Dan Dhaliwal for serving as editor on a previous version.

Submitted: November 2006
Accepted: January 2009
Published Online: September 2009

JEL Classifications: *M41; G14.*

Data Availability: *Data are available from public sources indicated in the text.*

I. INTRODUCTION

Demand for financial reporting and disclosure arises from information asymmetry and agency conflicts between managers, outside investors, and intermediaries. Disclosures and institutions facilitating credible disclosure between managers and investors play important roles in mitigating these problems.[1] Corporate disclosures, reports in the business press, and analysts' reports and discussion of corporate performance all enhance the information reflected in stock prices. That is, they reduce information asymmetry between the average investor and informed market participants, e.g., company management. The consensus among financial economists is that a rich disclosure environment and low information asymmetry have many desirable consequences. These include the efficient allocation of resources in an economy, capital market development, liquidity in the market, decreased cost of capital, lower return volatility, and high analyst forecast accuracy.[2] In this research, we study the consequences of disseminating financial information to the investment community through three sources, corporations, analysts, and business press, on the firm's capital market environment, i.e., the cost of capital, return variability, and analyst forecast error dispersion.

Evidence of the effect of disclosure on the firm's capital market environment, in particular, the cost of capital, is sparse (see Healy and Palepu [2001] for a review of the literature). Evidence is mixed in the few studies that examine the issue (e.g., Botosan 1997; Botosan and Plumlee 2002; Francis et al. 2006; Core et al. 2008). Conclusions from previous research on the effect of disclosures in the annual report (e.g., Botosan 1997) on the cost of capital are tenuous because disclosures analyzed in previous research are far from comprehensive and the measurement of disclosure proxies is generally subjective, not objective. With respect to studies that use objective measures (e.g., Francis et al. 2006), critics contend the proxies might not capture disclosure quality (see Liu and Wysocki 2006) and/or the accruals quality is not priced as a distinct factor beyond the Fama-French three factors (see Core et al. 2008).

We believe an important reason for the mixed evidence on a relation between disclosure and measures of a firm's capital market environment is that disclosure is assumed to have a unidirectional effect on the cost of capital or return volatility, or on other measures of a firm's capital market environment. While a unidirectional relation is expected using the *quality* of disclosure as the construct for disclosure, empirical measures of disclosure quality are likely influenced also by disclosure *content.* Thus, empirical measures may not be unambiguous proxies for disclosure quality. In this study, we move the literature forward by testing a directional link—whether disclosures containing favorable and uncertainty-reducing content result in lower cost of capital, and conversely disclosures containing unfavorable and uncertainty-increasing information or content cause the cost of capital to rise. We make similar predictions with respect to return volatility and analyst forecast error dispersion.

Ours is the first study to document systematic evidence of the capital market environment effects of disclosures culled from a virtually exhaustive set of sources from the print medium. Our analysis of disclosures is not only comprehensive, but also the construction of disclosure measures is quantitative and amenable to replication. Our study is also distinct from previous studies on the relation between disclosures and cost of capital in that (1) we test for the effects of favorable and unfavorable disclosures, i.e., content, whereas past research has typically focused on the cross-sectional association between the level of disclosure quality and the cost of capital; and (2) we predict and test a directional relation between disclosure content and cost of capital, return volatility, and analyst forecast dispersion. Specifically, we examine whether favorable and unfavorable disclosures in a quarter affect firms' cost of capital, return volatility, and forecast dispersion.

---

[1] See Healy and Palepu (2001), and Lambert et al. (2007).

[2] See Diamond and Verrecchia (1991), Verrecchia (2001), Healy and Palepu (2001), Easley et al. (2002), Easley and O'Hara (2004), Leuz and Wysocki (2008), Liu and Wysocki (2006), Bushman and Smith (2001), Core (2001), and Lambert et al. (2007).

In addition to analyzing the impact of all disclosures on the capital market environment as proxied for by the cost of capital, return volatility, and analyst forecast error dispersion, we study the impact of disclosures segregated by source—management, analysts, and news stories in business press. There are economic reasons to study how the source of disclosures affects the firm's capital market environment. Disclosures by management, analysts, and news reporters differ on at least two dimensions. First, incentives facing management, analysts, and news reporters might differentially shade the contents of their disclosures or reports. In particular, management and investment analysts can have strong incentives to optimistically skew disclosures, whereas news reporters' incentives to be optimistic in their reports are muted. The capital market is likely to weight disclosures according to the credibility of the source of the disclosure. That is, on average, in pricing stocks we expect the market to filter out the bias and deemphasize noisy disclosures by the management, analysts, and news reporters.

Second, news stories in the business press are likely to be timelier than analysts' reports. Analysts frequently repackage and re-transmit available information from corporate disclosures and business press news stories in presenting in-depth analysis in their own reports (see Lang and Lundholm 1996; Frankel et al. 2006). While such reports might be valuable to the market participants in their investment decisions, they may also have limited price impact or limited implications for the cost of capital or return volatility or analysts' forecasts.

Our empirical analysis tests for the impact of disclosure content on (1) cost of capital, where the cost of capital is estimated using the Fama-French three-factor model (see Fama and French 1993), (2) firms' stock return volatility, and (3) dispersion in analysts' forecasts, where the latter two measures are likely to be correlated with the firms' cost of capital. Our findings are based on a large content database of disclosure text constructed from disclosure content published on four electronically available information sources: Dow Jones Industrial, Investext, Factiva, and SEC EDGAR.

The content analysis classifies disclosures as favorable or unfavorable, with quantitative measures that indicate the intensity or degree of favorableness. The measures are based on an analysis of the content in each "story" (e.g., an analyst report).[3] We find that when content analysis indicates favorable disclosures, the firm's cost of capital, stock return volatility, and dispersion in analysts' earnings forecasts decline significantly. In contrast, unfavorable disclosures are accompanied by significant increases in the cost of capital, stock return volatility, and analysts' earnings forecast dispersion.

We next analyze disclosures by source: corporations, analysts, and the business press. First, we find that the market discounts the impact of management's statements. Positive news statements by management do not materially affect the firm's cost of capital, suggesting that management's favorable disclosures may not be credible. When management offers negative news disclosures, return volatility rises, and analyst forecast dispersion widens, but we do not find a cost of capital effect as predicted. This is consistent with bad news from management's mandatory disclosures not being timely or that the measures of cost of capital are not accurate.

---

[3] Content assessment measures would be favorable, for example, when disclosures contain references to earnings, cash flow, and sales increases or growth, favorable product reviews, enhanced management credibility, reduced uncertainty about the firm's products and markets, etc.

Second, both positive and negative news disclosures from analysts appear to be heavily discounted. Analysts on the whole are slightly less positive in their disclosures on firms, and similarly less negative (more favorable) in negative disclosure. The impact of their disclosure on cost of capital is insignificant. Intuitively, analysts have source credibility problems. They are seen as responding to market changes after they have taken place, discounting their impact.

Finally, we consider the impact of positive and negative news in business press. Here the evidence shows that both positive and negative news disclosures in the business press impact the cost of capital, return volatility, and analyst forecast dispersion. Positive news disclosure decreases all three, negative news disclosure increases it. Intuitively, the credibility and timeliness of news disclosure by the press are higher than that for management or analysts. The press is less affected by incentive concerns (agency), and this is reflected in the consistency of evidence.[4]

Section II presents the background and key arguments for our main hypothesis that the credibility and timeliness of the source of disclosures and the content of the disclosures affect the firm's cost of equity capital as well as volatility of returns and earnings (forecasts). Section III discusses our data and content analysis methodology. We detail our building of the disclosure text data set, our content analysis methodology, and our financial risk measures and firm characteristics. In Section IV we present our results. Our concluding section discusses the management implications of our findings and opportunities for future research.

II. BACKGROUND AND HYPOTHESIS DEVELOPMENT

In this section we provide the background for our main hypothesis that the credibility and timeliness of the source of disclosures affect their impact on firm's cost of equity capital, return volatility, and analyst forecasts, i.e., the firm's capital market environment. Economic rationale for the hypotheses predicting the directional effects of disclosure content is developed. The analysis is extended to motivate hypotheses about the differential effects of disclosures by corporations, analysts, and business press. We assume that disclosures in general can affect the firm's capital market environment, including the cost of capital and return volatility, which is extensively researched in the literature (see references in footnote 2 and discussion below).

Credibility and Timeliness of Various Sources of Disclosure

Credibility and Timeliness of Management Disclosures

We analyze SEC-mandated corporate disclosures, i.e., 10-K and 10-Q filings, which include the annual report. These disclosures contain a combination of (1) disclosures through the Management's Discussion and Analysis (MD&A) of the firm's performance that are designed to bring the market's expectations in line with the management's superior information; and (2) mandated other disclosures such as financial footnotes that are largely an affirmation of the information already available to the market participants, but could occasionally include qualitative and forward-looking information, e.g., discussion of loss contingencies and litigation risk.

---

[4] Bushee et al. (2008) find that press coverage around earnings announcements increases absolute returns and lowers information asymmetry (as inferred from bid-ask spreads). An increase in absolute returns suggests an increase in return variability associated with the arrival of information, whereas decreased spreads indicate press reports reducing the asymmetry of information among investors. Their results complement our work, although they do not content analyze the press reports to classify them as carrying favorable or unfavorable news.

In utilizing corporate disclosures, market participants are likely to recognize that management has an incentive to portray a self-serving assessment of past corporate performance and future outlook. Management's incentive to favorably skew disclosures might stem from the use of accounting information by current and potential creditors in debt contract negotiation, by boards of directors in determining management compensation, by the labor market for executives in assessing the quality of the management, by labor unions in wage and labor contract renegotiations, and by capital market participants in setting stock prices and in determining the cost of capital.[5] In light of these incentives, management's favorable disclosures may not be credible to market participants. Of course, in making any favorable, optimistic disclosures, the management runs the risk of being accused of making misleading and fraudulent disclosures, i.e., face litigation risk. Litigation risk thus serves to discipline managers from making overly optimistic disclosures. However, the disciplining forces might be less operative when disclosures are qualitative and long-term in nature (e.g., discussion in the MD&A section) rather than quantitative (e.g., point estimate for an EPS forecast) and short-term (e.g., next quarter's sales).

In contrast to good news disclosures, management has an aversion to disclose bad news. So, when it volunteers such information, it is believable. That is, management's unfavorable voluntary disclosures are more credible regardless of whether they are qualitative or quantitative.

While managements' bad news disclosures are likely to be credible, there are conflicting forces affecting the timeliness of management's unfavorable disclosures. Litigation risk facing the firm under the Securities and Exchange Commission's Rule 10b-5 (Skinner 1994, 1997) is likely to motivate the manager to disclose bad news early. The Rule together with other securities laws state that management has a duty to disclose material information as and when it becomes known to the management. The rationale is that new information might render the management's previous disclosures misleading. Alternatively, in light of the new information, the stock would be misvalued in the absence of disclosures to correct the mispricing. However, the securities laws are likely to have an asymmetric impact on corporate disclosure practices with respect to good and bad news disclosures. Litigation against a firm is more successful if the management is accused of withholding bad news such that the stock price remained above its fundamental value. Consistent with this rationale, Kellogg (1984) finds that buyers' claims of losses are 13 times as likely as sellers' claims of forgone profits, i.e., an opportunity loss. Therefore, management has an incentive to volunteer bad news on a timelier basis than good news (see Kasznik and Lev 1995; Skinner 1994, 1997).

However, management's career concerns might motivate managers to delay the release of bad news (see Kothari et al. 2009). Career concerns include management's concern about the effect of reporting adverse financial performance on promotion, outside employment opportunities, termination, and post-retirement opportunities. The manager might attempt to mitigate these concerns by delaying the release of bad news.

---

[5] See Watts and Zimmerman (1986) and Fields et al. (2001) for various motivations for management optimism in disclosures and for summaries of related evidence.

In light of the conflicting incentives to accelerate and delay the release of bad news, actual practice becomes an empirical issue. Kothari et al. (2009) find that managers successfully withhold bad news before making voluntary disclosures. As a result, when managers do disclose the bad news in voluntary management forecasts, it is, on average, of a substantial magnitude, i.e., generates a large, negative price response. In contrast, managers disclose or leak good news early. We do not examine management's voluntary forecasts, but limit our analysis to mandated, periodic management's discussion of firm performance and outlook for the future in the annual 10-Ks and quarterly 10-Qs. Preceding arguments and evidence suggest that management disclosures in these reports is likely to be more credible if it communicates bad news, but the timeliness of good and bad news is difficult to predict.

Credibility and Timeliness of Analyst Reports

Analysts are important information intermediaries in capital markets. They expend real resources on gathering information and producing research reports. These activities are typically assumed to enhance the ability of prices to reflect available information. However, several forces influence the credibility and timeliness of analyst research, which shape their price impact and effect on cost of capital.

Strong economic incentives stemming from brokerage commissions and investment-banking revenues potentially taint the quality of analyst research. United States Senator Joseph Lieberman claims that the millions of dollars analysts earn from investment-banking services "compromise analysts' objectivity" and therefore "the average investor should take their bottom line recommendations with at least a grain of salt, if not a whole bucket," (see Hamburger and Burns 2002). Academic evidence is mixed in suggesting that analysts compromise their objectivity and issue optimistic forecasts and stock recommendations as a result of their economic incentives arising from brokerage revenues and investment-banking relationships.[6] Questions surrounding the objectivity and bias in research together weaken the credibility of analyst research, which, in turn, would diminish its influence on the pricing or on the cost of capital.

With respect to the timeliness of information in analyst reports, previous research offers compelling evidence that analyst reports do convey new information that affects security prices (see Lys and Sohn 1990; Frankel et al. 2006). However, the magnitude of price impact is typically small. This might be because considerable analyst research is simply a retrospective analysis of past events, which repackages and retransmits publicly available information in corporate disclosures and news stories (Lang and Lundholm 1996; Beaver 1998, 10; Frankel et al. 2006). If this activity dominates the contents of analyst research, it might not be particularly timely in providing much new information to the market. Worse, evidence in Lys and Sohn (1990) and Guay et al. (2006) suggests that analyst forecasts are sluggish in that they fail to fully incorporate the information available as reflected in stock prices at the time of their forecasts.

In summary, questionable credibility and the absence of timely forecasts and research all make it less likely that analysts' research reports would contain much information.

---

[6] See Lin and McNichols (1998), Michaely and Womack (1999), Dechow et al. (2000), Bradshaw et al. (2003), O'Brien et al. (2005), Kolasinski and Kothari (2008), Clarke et al. (2007), and Ljungqvist et al. (2007).

Credibility and Timeliness of Business Press News Stories

We expect news stories in the business press to be quite timely with an emphasis on factual reporting. News reporters and the business press typically do not have strong economic ties and relationships with individual firms. Moreover, each reporter reports on a large number of corporations as new developments take place with individual firms. Therefore, we expect news stories to have both credibility and timeliness and, hence, the contents of news stories are likely to be informative about the cost of capital.

Hypothesis Development: Effects of Disclosure Content on the Cost of Capital, Volatility, and Earnings Forecasts

We begin this section by summarizing the effects of disclosure on a firm's capital market environment. This discussion focuses on disclosure per se, not the content, i.e., good or bad news. We then develop a directional link between the content of disclosures and the capital market environment. Specifically, we explain why good (bad) news might presage a decline (an increase) in the cost of capital, return volatility, and analyst forecast dispersion.

Longstanding theoretical literature suggests a link between disclosure and the equity cost of capital. Disclosure's impact on the cost of capital is hypothesized in at least three ways in the literature. First, disclosure about a firm is likely to enhance the liquidity of a security because disclosures tend to reduce the degree of information asymmetry between insiders and outsiders of a firm and also among various groups of current and potential investors of a firm (e.g., Diamond and Verrecchia 1991; Amihud and Mendelson 1986; Easley and O'Hara 2004). Information asymmetry widens the adverse-selection component of the bid-ask spread demanded by the market makers and thus increases the cost of trading in a security. In equilibrium, security prices settle at levels that yield investors equal rates of risk-adjusted return, net of the transaction cost, on each security.

Second, investors do not know the true values of the parameters of expected rates of return on securities. Literature on estimation risk demonstrates that errors in estimating the parameter values represent a source of non-diversifiable risk and, therefore, constitutes a component of the cost of capital (e.g., Klein and Bawa 1976; Barry and Brown 1984). Disclosures (i.e., information) enable investors to reduce the degree of error in estimating the parameters, which lowers a firm's cost of capital.

Third, Lambert et al. (2007) theorize that the covariance of a firm's cash flows with other firms' cash flows decreases in the precision of disclosures. This reduces the cost of capital of the firm making more precise disclosures. An additional reason for the effect of disclosures on a firm's cost of capital and volatility stems from the improvement in the investment decisions, which results from the enhanced quality of information available to managers. The improvement is in part a result of lesser uncertainty about cash flows as a consequence of more precise information.

For all of the above reasons, *ceteris paribus* disclosure levels are expected to be negatively associated with the cost of capital and volatility. In this study we assume all of the reasons as contributing to the negative relation without attempting to disentangle the relative importance of each. However, preceding discussion suggests a unidirectional link between disclosure and cost of capital and volatility. That is, the greater the amount of disclosure, the lower the cost of capital and volatility. We move the literature forward by testing a directional link—whether disclosures containing favorable and uncertainty-reducing content lower a firm's cost of capital and return volatility, and conversely disclosures containing unfavorable and uncertainty-increasing information or content cause the cost of capital and volatility to rise. Similar predictions apply with respect to the dispersion in earnings forecasts. Our empirical predictions are grounded in prior theoretical work and survey data discussed below.

In an efficient market, price change associated with news is either the effect of news about expected future cash flows or news about the risk of the cash flows, i.e., discount rates. Thus, even if information about cash flows is assumed to be unrelated to information about the risk of those cash flows, on average, unfavorable information would be associated with an increase in cash flow risk and, conversely, favorable information would be associated with a decline in the risk of expected future cash flows. French et al. (1987) develop and test this hypothesis over short horizons and Chan (1988) and Ball and Kothari (1989) test it over long horizons. Chan (1988) and Ball and Kothari (1989) conduct their analysis within the CAPM framework. Ball and Kothari (1989) find that both changes in the CAPM cost of capital and return volatility are negatively related to news content and therefore price changes.

A second reason for predicting a negative relation between disclosure content and the cost of capital as well as return volatility is the effect of news on the firm's leverage (Christie 1982; Galai and Masulis 1976; Ball and Kothari 1989). Price changes (or news) are negatively related to discount rates via the leverage effect. In a CAPM framework, *ceteris paribus*, price changes correlate negatively with leverage, and leverage and expected returns are positively related (see Galai and Masulis 1976). To the extent firms do not dynamically and immediately adjust financial leverage following news affecting the firm's equity value, we expect a negative relation between news content and leverage. This, in turn, predicts a negative link between news content and cost of capital changes and volatility changes.

A further stream of literature posits a negative relation between disclosure content and cost of capital as a result of the effect of disclosure content on the degree of information asymmetry between the firm (management) and the investment community. The literature has origins in Glosten and Milgrom (1985), Diamond and Verrecchia (1991), Verrecchia (2001), Easley et al. (2002), Easley and O'Hara (2004), and the literature on limited information risk, parameter uncertainty, and liquidity risk.[7] While much of the literature focuses on the consequences of the precision of information, i.e., the variance effect, recent studies, including Ng et al. (2009), develop a link between disclosure content and changes in cost of capital and return volatility within the framework of information asymmetry between the managers of a firm and the investment community. They argue that because news content is likely to have a directional effect on the uncertainty with which investors can forecast a firm's future cash flows, the effect of disclosure content on the adverse-selection component of the bid-ask spread is likely to be directional. That is, good news is likely to reduce the adverse-selection component of the bid-ask spread, whereas bad news is expected to increase it. Specifically, they argue that performance reports, i.e., disclosure content, predict volatility in future earnings: measures of good performance are associated with lower volatility in future earnings, while measures of poor performance are associated with higher volatility in future earnings. Market makers use performance reports in assessing both the expected value of the firm and uncertainty about that value. Therefore, performance reports affect adverse selection costs through market makers' assessments of uncertainty. The effects ultimately are predicted to be manifested in the firm's cost of capital, return volatility, and variability of the analysts' earnings forecasts.

---

[7] See Klein and Bawa (1976), Barry and Brown (1984), Chordia et al. (2001), and Pastor and Stambaugh (2003).

Since we content-analyze a large number of textual disclosures, unlike previous research, we are in a position to quantify the favorable or unfavorable nature of disclosure, not simply the amount of disclosure. The ability of content analysis to distinguish favorable, risk-reducing disclosures from unfavorable, risk-increasing disclosures permits us to hypothesize that the former would improve the firm's capital market environment, i.e., lower the cost of capital, return volatility, and earnings forecast dispersion, whereas the latter would worsen the capital market environment.

Business executives corroborate the predicted directional link between disclosure content and risk effects. Survey evidence in Graham et al. (2005) finds that "executives believe that less predictable earnings—as reflected in a missed earnings target or volatile earnings—command a risk premium in the market." Such a belief among corporate executives indicates that disclosure of unfavorable news would prompt investors to demand a higher return on the company's stock to compensate for the increased risk, i.e., cost of capital, return volatility, and earnings volatility.

III. DATA AND CONTENT ANALYSIS METHODOLOGY

In this section we discuss disclosure data obtained from electronic information sources, which we used in conducting content analysis, and then describe data on financial risk measures and firm characteristics.

Disclosure Text Data

The most widely available source of corporate disclosure is financial statements. These are SEC-mandated disclosures for all U.S. publicly traded corporations. In addition to financial statements, managers also provide other information to investors. Some, such as footnotes and MD&A, are disclosed within the financial report. Other disclosures are provided voluntarily through other information channels, including analyst presentations and conference calls, press releases, and Internet sites. Through these disclosures, managers provide information that facilitates external users of financial reports to better understand the true economic picture of the business.

Disclosures are accompanied by third-party certification to help assure credibility. That is, in the absence of third-party certification, there would be questions in investors' minds about the credibility of the information contained in the disclosures. For disclosure in financial statements, credibility is typically enhanced in two ways. First, statements are prepared using generally accepted accounting standards and conventions. Standards can potentially reduce processing costs for outsiders and curb opportunistic reporting by managers. Second, independent audit firms certify whether managers' financial reporting decisions are consistent with accounting standards and conventions.

Importantly, management disclosures made outside of financial statements are not directly certified by independent parties. Rather, their source credibility is enhanced through their use and evaluation by outside parties, including financial analysts and business journalists. These "information intermediaries" provide a measure of implicit certification of management's financial statements and nonfinancial disclosures. However, this certification is not in any way formal nor can it be argued to be necessarily complete. Our analysis is designed to empirically assess whether the credibility issues with respect to disclosures from various sources, as discussed above, are manifested in their impact on the firms' cost of capital.

We obtain data for our study from four electronic data information sources: Dow Jones Interactive (Dow Jones, Inc.), Investext Plus (Infotrac), Factiva (Dow Jones & Reuters), and EDGAR (Securities and Exchange Commission, U.S. Government). Text-retrieval software was developed and utilized to efficiently download disclosure texts in their entirety for a company research sample frame of 889 firms over the time period 1996–2001 (Riloff and Hollaar 1996).

We downloaded all disclosure texts available, by company by year, from three sources: (1) corporate reports (EDGAR), (2) analyst disclosure and briefings (Investext and Factiva), and (3) disclosures made by the general business press (more than 400 content sources available in Dow Jones Interactive and Factiva). All disclosure texts were coded by company ID, date and source of publication, and, when available, by author. The resulting content data set is very large. For the 889 companies over the six-year period, 326,357 texts were downloaded, representing over a million pages of source material. The 889-company sample was drawn from four industry sectors: Financial Services (225 companies), Technology (197 companies), Pharmaceutical (72 companies), and Telecommunications (395 companies), and the data was coded to enable industry sector analysis. These are all the companies in these four industries that also had machine-readable financial statement and stock return data on Compustat and CRSP.

Finally, source material was organized into business quarters, corresponding to the quarterly financial reporting cycle. Each company in the data set therefore had 72 observations: disclosure content aggregated into 24 time periods (6 years × 4 quarters per year), by three sources of disclosure (corporate, analyst, business press). The list of companies in the data set and the business publications included in the Dow Jones and Factiva electronic data services are available upon request.

Classification of Text into Content Categories

Our first step in preparing the data for content analysis was to classify each disclosure text using a business word classification scheme that we developed. Words associated with six-categories of the business environment, strategy, operations, human resources, and others were defined and used to build a classification scheme using Riloff (1993) as a guide. The six categories identified words and word phrases associated with:

1. market risk, industry structure, and competitive forces, and the external market and/or external environment of the firm;
2. firm risk, the development, and execution of firm strategy and strategic intent;
3. organizational risk, the building of organizational capital, and human resource management;
4. reputational risk, the image, brand, and reputation of the company;
5. performance risk, the investment, and financial performance of the firm; and
6. regulatory risk, the announcement, and impact of governmental regulation on the company.

A classification scheme and the words and word phrases making up each category are available upon request.

All disclosure text was filtered through the classification scheme, producing six subsample data sets containing words and word phrases matching the words and word phrases defined in the six categories. Our subsequent analysis focused on content analyzing within and across the subsample data sets and the (original) master text data set.

Content Analysis

The underlying principle in content analysis is that the many words of a text can be classified into many fewer content categories, where each category consists of one or many similar words or word phrases, and that each word or phrase occurrence can be counted and the counts compared analytically. Word and phrase similarities are based on the precise meanings of the words themselves (for example, grouping synonyms together), or may be based on word groupings sharing similar connotations (for example, grouping together several words associated with a concept such as market share, revenue growth, or forecasted earnings). Content analysis therefore, can be useful in analyzing different types or "levels" of communication, as defined by the meaning of the words themselves.

A flowchart of the typical process of content analysis research is presented in Figure 1. The study is first specified theoretically and then conceptually in terms of an information model, content (data samples), and the hypotheses to be addressed. The study is then operationalized with respect to the variables (content categories) and coding schema to be used for classification of words and word groups for analysis. Third, the content texts are analyzed using content analysis procedures and the results interpreted.

Analysis Procedures and Measurement in Content Analysis

Figure 2 illustrates the specific sequence of steps used in our content analysis. We first passed the complete, 889-company disclosure content data set through a business dictionary of terms we defined to cluster disclosure texts with similar words and word meanings together. The clustering was not exclusive, that is, a disclosure text could be clustered in one category, in more than one, or in no category. But within each business category, all disclosure texts were matched with words and word meanings we defined in the category.

FIGURE 1
Flow of Content Analysis

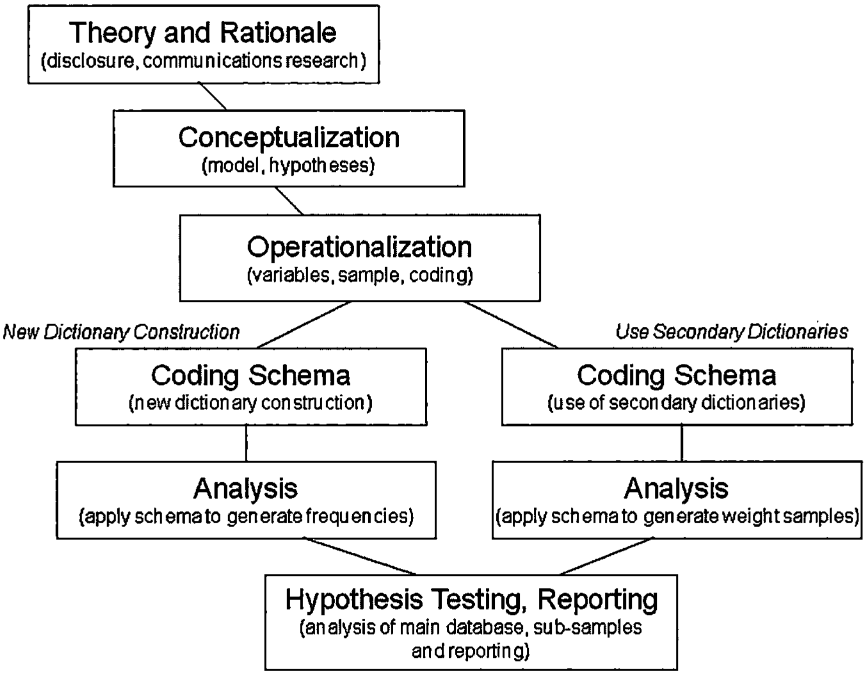
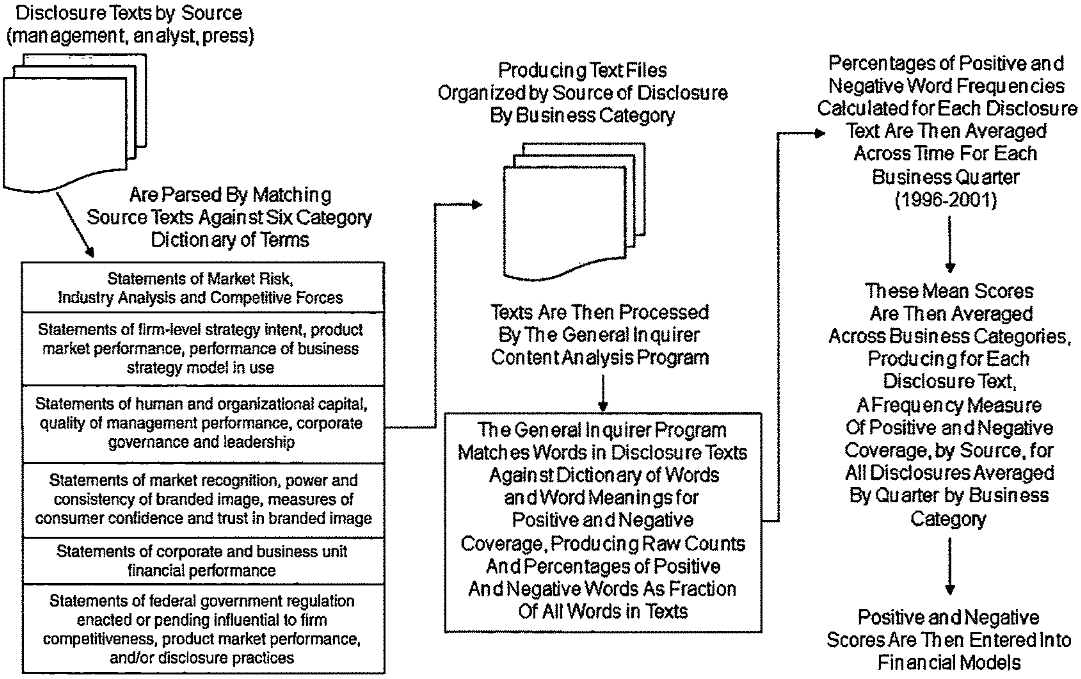

Theory and Rationale
(disclosure, communications research)
Conceptualization
(model, hypotheses)
Operationalization
(variables, sample, coding)
New Dictionary Construction
Use Secondary Dictionaries
Coding Schema
(new dictionary construction)
Coding Schema
(use of secondary dictionaries)
Analysis
(apply schema to generate frequencies)
Analysis
(apply schema to generate weight samples)
Hypothesis Testing, Reporting
(analysis of main database, sub-samples and reporting)

FIGURE 2
Analysis Steps in Constructing Measures of Positive and Negative Coverage of Each Company by Source of Disclosure


Disclosure Texts by Source (management, analyst, press)
Are Parsed By Matching Source Texts Against Six Category Dictionary of Terms
Statements of Market Risk, Industry Analysis and Competitive Forces
Statements of firm-level strategy intent, product market performance, performance of business strategy model in use
Statements of human and organizational capital, quality of management performance, corporate governance and leadership
Statements of market recognition, power and consistency of branded image, measures of consumer confidence and trust in branded image
Statements of corporate and business unit financial performance
Statements of federal government regulation enacted or pending influential to firm competitiveness, product market performance, and/or disclosure practices
Producing Text Files Organized by Source of Disclosure By Business Category
Texts Are Then Processed By The General Inquirer Content Analysis Program
The General Inquirer Program Matches Words in Disclosure Texts Against Dictionary of Words and Word Meanings for Positive and Negative Coverage, Producing Raw Counts And Percentages of Positive And Negative Words As Fraction Of All Words in Texts
Percentages of Positive and Negative Word Frequencies Calculated for Each Disclosure Text Are Then Averaged Across Time For Each Business Quarter (1996-2001)
These Mean Scores Are Then Averaged Across Business Categories, Producing for Each Disclosure Text, A Frequency Measure Of Positive and Negative Coverage, by Source, for All Disclosures Averaged By Quarter by Business Category
Positive and Negative Scores Are Then Entered Into Financial Models

We then submitted each clustered subset of disclosure texts, and the original (complete) content data set, to the General Inquirer (GI) content-analysis software. The GI is a text-mapping tool. Its software algorithm counts words and word meanings present in submitted texts against an exhaustive set of dictionary-supplied categories. As we were especially concerned with the degree and strength of positive and negative statements in disclosure texts, we utilized two GI dictionary categories, "Positive" and "Negative," which are categories of several thousand words and word meanings associated with positive and negative outlook. Using GI to accurately gauge the frequency and strength of positives and negatives in the disclosure texts submitted would therefore tell us much about the consistency of views about the firm and its risks and promise in the marketplace.

Once the GI software returned for each disclosure text the raw counts and word frequencies (percentage of matched words against the number of all words in the text), we then calculated the average score by quarter for each quarter in the sample time period of 1996–2001. Finally, for each calculated mean score, we then calculated the mean score for the six business categories we defined. The net result was that for each disclosure text in the sample, we calculated a mean score for strength of positive and negative news coverage, by source, by business quarter, and used that score for input into financial analysis.

To illustrate a specific example of how our content procedures worked in practice, and the form of the output matrix calculated, we briefly describe the analysis of two example disclosure texts published in the *Wall Street Journal* on Amgen Corporation. The two texts (both business news stories), published in February and April 1996, report positive and negative news on Amgen. On February 1, 1996, the *Journal* published a story reporting that clinical research undertaken on the effects of Leptin, the company's experimental obesity drug, showed that the drug did not have the hoped-for benefits in a controlled patient study. The second story, published on April 18, 1996, reported the company's 32 percent increase in net quarterly income and the filing of an application for testing a new drug designed to combat Hepatitis C. The text of both stories is available upon request.

We submitted the two texts to the General Inquirer (GI) content-analysis package. Developed and written by Professor Philip Stone and colleagues at Harvard University, and later extended by Vanja Buvac and colleagues, the General Inquirer (GI) software program is essentially a tool for mapping content.[8] The GI software "maps" (matches) words in each text file with words and word senses in dictionary-supplied categories. That is, the software algorithm identifies and counts word frequencies in the submitted text files, matching them against words and/or word meanings present in one or more dictionary of categories. The currently distributed version of GI combines the "Harvard IV-4" dictionary content-analysis categories, the dictionary content-analysis categories defined by Lasswell, and five categories based on the social cognition work of Semin and Fiedler.[9] Each category contains a list of words and word senses. For example, the category "negative" is the largest, with 2,291 entries of words or word phrases denoting "negative."[10] When text is submitted for processing, the GI matches words and word phrases with those in its dictionaries, and separates inflected word forms. For example, the root word "sell" would match "selling" if selling was not a separate category entry, and the word sense "sell cycle," for example, would be correctly identified by software routines designed to disambiguate word senses. That is, "cycle" would be correctly identified as specifying the time periodicity of selling, and not a word short for bicycle or motorcycle. These English-stemming procedures, integrated with English dictionaries and routines for separating English word senses, comprise the software's extensive dictionaries and disambiguation routines. The complexity of the dictionary and disambiguation routines limits the current GI software to English text applications.[11]

The results of our simple, two-disclosure text submission to the software appear in a 4 × 6 matrix. In interpreting the output, for each text, there are two rows of data. The first corresponds to the raw frequency count of matched words ("r"), and the second gives the percentage of matched words over the total number of words in the text. The February 2, 1996 *Wall Street Journal* article, for example, contains 635 words, of which 20 are positive and 21 negative. Positives comprised 3.1 percent of the total number of words in the text, and negatives slightly higher, 3.3 percent. There are 116 words that do not match any of the words supplied in the GI dictionary categories, or 18.3 percent. In the case of the April 18th *Wall Street Journal* article, the corresponding numbers are the text contains 423 words, 19 positives and 8 negatives, or 4.4 percent of the words in the disclosure are positive and 1.8 percent are negative. Seventy-seven words, or 18.2 percent, are leftovers (unmatched with GI dictionary categories).

---

[8] The General Inquirer's home page is at http://www.wjh.harvard.edu/~inquirer/.

[9] See Semin and Fiedler (1992, 1996). Categories are described on the GI home page, http://www.wjh.harvard.edu/~inquirer/.

[10] Example entries for the positive and negative categories are available upon request.

[11] The GI package is designed to pick up negative qualifiers of positive sounding terms like growth as in "sales are not expected to grow" and characterize such text as negative, not positive.

Interpreting results in content analysis is always a question of comparing word counts and word frequencies across texts and categories. In this example, the first article is much more negative than the second article, and, as well, less positive. Conversely, the second article is much more strongly positive and significantly less negative in its word usage and meaning.

In analyzing the large number of disclosure texts in the database, our procedure calculates the number of positives and negatives in each disclosure text, for each of the six business categories defined (market, firm strategy, etc.), by business quarter, by source. The resulting matrix of counts and percentages calculated for each disclosure text is then used as input into our econometric analysis.

Financial Risk Measures and Firm Characteristics

We use three measures of risk to assess the impact of disclosures on firm risk: cost of equity capital, standard deviation of stock returns, and standard deviation of the analyst forecast errors. We describe each measure below.

We estimate the equity cost of capital using the Fama and French (1993) three-factor model, where the size factor is defined as small minus large firm returns (SML), the book-to-market factor is defined as high minus low book-to-market firm returns (HML), and the market factor is defined as the excess return on the CRSP value weight portfolio ($R_m - R_f$).[12] We obtain monthly time-series returns on the three factors, SML, HML, $R_m - R_f$, from Kenneth French's website. The loadings on the factors, b, s, and h, are slope coefficients estimated from the following regression model for firm i:

$$R_i - R_f = a_i + b_i\,[R_m - R_f] + s_i\,SML + h_i\,HML + e_i. \qquad (1)$$

We reestimate the three-factor model each quarter for each firm using a rolling window of five years of monthly returns ending in the quarter under examination. Firm i's estimated loadings, i.e., estimated b, s, and h coefficients, multiplied by the average returns for the three factors from 1963–2000 gives the cost of capital for firm i (see Fama and French 1993). We then annualize the number, which is our cost-of-capital measure.[13]

Standard deviation of stock returns is a commonly used risk measure. The greater the uncertainty of cash flows generated by a firm, i.e., the risk of equity, the greater is the standard deviation of returns. This measure is also an indication of the infrequency of information reaching the market and the degree of information asymmetry among market participants. If some individuals are informed and others are less informed, then the trading in the stock is less frequent and stock prices do not reflect a rich body of information. Under these circumstances, disclosures by a corporation or analysts or disclosures in business press can all influence the degree of uncertainty. Favorable disclosures can inform the market of higher levels of cash flows than previously expected and/or reduce the uncertainty in the market. We measure the standard deviation of returns using daily return data for each firm every quarter.

---

[12] While there is no debate about what should be the measure of return volatility, conversely there is no agreement among researchers about the measure of cost of capital. Our tests for the impact of disclosures on the firms' capital market environment assume that the cost of capital effect is manifested in the firms' sensitivity to the Fama-French three factors. This is consistent with the analysis in Lambert et al. (2007). That is, the cost of capital effects associated with disclosures do not stem from a distinct "information risk" factor (see Francis et al. 2005; Core et al. 2008; Liu and Wysocki 2006, for a debate on whether accounting quality is a distinct information risk factor). To the extent that the disclosures' effect on cost of capital is not only subsumed by the Fama-French three factors, but also produces a distinct information risk effect, our estimated cost of capital effects are likely to be understated.

[13] An alternative to using the Fama-French three-factor model to estimating the cost of capital is an implied cost of capital estimate (e.g., Claus and Thomas 2001; Gebhardt et al. 2001; Easton 2004; Ohlson and Juettner-Nauroth 2005; Guay et al. 2006). Such an approach is straightforward to implement, but in our opinion it is not the most appropriate. Specifically, we do not follow this approach because our goal is to estimate a change in the cost of capital estimate associated with the disclosure news, favorable or unfavorable. Unless analysts' long-term earnings growth forecasts change over the quarterly horizon, we will not be able to estimate a change in the cost of capital estimate. However, analysts' long-term growth forecasts are relatively sticky and change only in response to fairly extreme news (see Guay et al. 2006).

Standard deviation of analyst forecast errors indicates the extent to which the market is surprised by earnings announcements. Favorable disclosures are expected to reduce uncertainty and enable analysts to better forecast earnings, which would, in turn, reduce the magnitude of their forecast errors, i.e., the standard deviation of forecast errors.

While our goal in the research is to test for the effect of disclosures on the firm's cost of capital, we control for the effect of other determinants of the cost of capital to avoid drawing erroneous inferences. We consider three firm characteristics that previous research shows as significant determinants of the cost of capital, standard deviation of returns, and analyst forecast errors: firm size, book-to-market ratio, and leverage. Small firms are more risky than large firms in part because small firms typically are a single-project firm, i.e., a firm with undiversified portfolio of assets and projects. We measure firm size as the market capitalization of a firm's equity capital, i.e., number of shares outstanding times the share price, at the beginning of each quarter.

Book-to-market ratio increases in risk in part because the market valuation of equity is in the denominator of the ratio. Successful firms with expectations of a steady stream of high levels of future cash flows are highly valued in the market, which drives their book-to-market ratio down. These are typically considered to be low-risk firms. In contrast, if the market has little confidence in a firm and thus perceives the cash flow stream to be uncertain and not too high, then the market capitalization of such a firm would be low. This drives up the book-to-market ratio, so high book-to-market ratio proxies for high-risk firms. Book-to-market ratio is calculated as a ratio of the book equity of a firm divided by its firm size, i.e., the market capitalization of equity.

Finally, risk increases in the firm's financial leverage, measured as leverage as the ratio of long-term debt to the total assets of the firm.

IV. RESULTS

This section summarizes the empirical effects of disclosure on the three measures of risk: cost of capital, return volatility, and dispersion of analyst forecast errors. We begin with descriptive statistics and correlations among the variables. The main results describe regressions of various risk measures on content-analysis-based measures of disclosure by source.

Descriptive Statistics

Tables 1 and 2 provide univariate descriptive statistics and simple bivariate correlations for our dependent and independent variables, by industry sector and by all sectors. The number of observations reported in each panel represents the number of firms times the number of quarters for which sufficient data are available to compute the variables. Some firm-quarters are excluded from analysis based on insufficient data. The average cost of capital for the sample firms is 14.8 percent, with inter-quartile range of 9.7 percent to 18.3 percent. The two rows corresponding to index variables *Favorable* and *Unfavorable* are constructed as the mean of the content-analysis based Positive and Negative measures for disclosures across all six business categories and disclosure source. The disclosure content on average is dominated by favorable rather than unfavorable news. The mean score

TABLE 1
Descriptive Statistics for Variables: Quarterly Data for 889 Firms from 1996 to 2001[a]

Panel A: All Industries, 5,350 Observations

	Mean	Std. Dev.	Min.	Q1	Median	Q3	Max.
Cost of Cap.	0.148	0.076	0.001	0.097	0.138	0.183	0.499
$\sigma(R)$	0.027	0.016	0.001	0.016	0.023	0.033	0.218
$\sigma(FE)$	0.948	1.938	0.000	0.065	0.206	0.780	12.943
Favorable	4.676	1.748	0.000	3.434	4.970	5.991	10.159
Unfavorable	1.674	0.712	0.000	1.188	1.629	2.146	5.139
Size	6.762	2.524	0.189	4.835	6.489	8.533	13.230
B/M	0.742	0.958	0.001	0.256	0.510	0.850	9.893
Leverage	0.163	0.176	0.000	0.017	0.101	0.256	0.837

Panel B: Pharmaceutical Firms, 962 Observations

	Mean	Std. Dev.	Min.	Q1	Median	Q3	Max.
Cost of Cap.	0.146	0.110	0.002	0.064	0.116	0.204	0.499
$\sigma(R)$	0.036	0.021	0.009	0.021	0.030	0.045	0.218
$\sigma(FE)$	0.842	1.369	0.029	0.049	0.169	1.212	6.130
Favorable	4.634	1.515	0.000	3.816	4.763	5.823	8.444
Unfavorable	1.907	0.746	0.000	1.421	1.953	2.456	4.322
Size	7.100	2.678	2.046	4.800	7.055	8.949	12.578
B/M	0.371	0.778	0.001	0.098	0.196	0.345	9.416
Leverage	0.107	0.142	0.000	0.000	0.049	0.172	0.729

Panel C: Telecom Firms, 987 Observations

	Mean	Std. Dev.	Min.	Q1	Median	Q3	Max.
Cost of Cap.	0.140	0.073	0.012	0.091	0.129	0.169	0.472
$\sigma(R)$	0.031	0.014	0.005	0.020	0.028	0.038	0.086
$\sigma(FE)$	2.364	3.299	0.000	0.206	0.911	4.107	12.943
Favorable	3.643	1.541	0.000	2.270	3.744	4.763	7.409
Unfavorable	1.544	0.720	0.000	0.933	1.540	2.039	3.633
Size	8.211	2.146	2.263	6.923	8.224	9.875	13.134
B/M	1.203	1.702	0.007	0.238	0.505	1.422	9.893
Leverage	0.216	0.183	0.000	0.071	0.200	0.300	0.837

Panel D: Financial Firms, 3,007 Observations

	Mean	Std. Dev.	Min.	Q1	Median	Q3	Max.
Cost of Cap.	0.152	0.063	0.001	0.109	0.144	0.183	0.426
$\sigma(R)$	0.022	0.012	0.001	0.014	0.020	0.027	0.175
$\sigma(FE)$	0.371	0.804	0.000	0.048	0.119	0.340	5.589
Favorable	4.983	1.800	0.000	3.696	5.356	6.320	10.159
Unfavorable	1.589	0.658	0.000	1.154	1.553	1.961	3.881
Size	6.135	2.300	1.683	4.429	5.906	7.281	12.467
B/M	0.716	0.567	0.045	0.411	0.608	0.873	8.673
Leverage	0.169	0.182	0.000	0.022	0.103	0.265	0.776

(continued on next page)

TABLE 1 (continued)

Panel E: Tech Firms, 394 Observations

	Mean	Std. Dev.	Min.	Q1	Median	Q3	Max.
Cost of Cap.	0.141	0.077	0.003	0.085	0.130	0.187	0.370
$\sigma(R)$	0.032	0.018	0.004	0.019	0.027	0.042	0.099
$\sigma(FE)$	1.560	1.819	0.110	0.164	0.970	1.654	6.233
Favorable	5.021	1.303	0.000	4.139	5.381	6.035	7.915
Unfavorable	2.077	0.727	0.000	1.530	2.077	2.583	5.139
Size	7.090	2.905	0.189	5.065	7.099	9.555	13.230
B/M	0.689	0.642	0.030	0.174	0.492	1.008	3.980
Leverage	0.121	0.136	0.000	0.002	0.084	0.203	0.723

---

[a] This table provides descriptive statistics for variables used in subsequent tests. To be included in the sample, each firm has sufficient data to compute the variables below. The time period is from 1996 to 2001.

Variable Definitions:

Cost of Cap. = expected annual cost of capital calculated based on Fama-French three-factor model over the past five years' return data;
$\sigma(R)$ = standard deviation of daily stock return over the contemporaneous quarter;
$\sigma(FE)$ = standard deviation of all available quarterly analyst forecast errors from 1996 to 2001;
Favorable = favorable assessment index, constructed as mean of the content-analysis-based scaled favorable content measures for six categories of disclosures across three sources, i.e., corporations, analysts, and business press;
Unfavorable = unfavorable assessment index, constructed as mean of the content-analysis-based scaled unfavorable content measures for six categories of disclosures across three sources, i.e., corporations, analysts, and business press;
Size = natural logarithm of a firm's market capitalization of equity at the beginning of the quarter;
B/M = book equity divided by market equity at the beginning of the quarter; and
Leverage = long-term debt divided by total assets.

for *Favorable* variable is 4.676 compared to 1.674 for *Unfavorable.* Since the period 1996 to 2001 was largely characterized by expansion in the U.S., dominance of favorable news is not surprising.

In Table 2, with the exception of the lack of a significant correlation between *Cost of Capital* and *Unfavorable,* all correlations are significant at the 0.01 or 0.05 level for all variables—that is, *Favorable* and *Unfavorable* assessments are strongly correlated with all financial measures studied, cost of capital, standard deviation of daily stock return, the firm's market capitalization of equity at the beginning of the quarter (natural logarithm), book equity divided by market equity at the beginning of the quarter, and leverage (long-term debt divided by the total asset). That all cross-correlations (minus one) are significant between content measures and financial measures suggests a strong and persistent association between the two.

Regression Results

Tables 3 and 4 report the results from multivariate regression models on composite news measures and source of news. These are Fama-MacBeth regressions estimated cross-sectionally in each quarter from Q1 1996 to Q4 2001. We report mean values of the time-series of estimated slope coefficients with Newey-West corrected time-series t-statistics. In addition, across all of the panels we report the average adjusted R^2. The panels report regression results separately on three dependent variables: cost of capital, the standard deviation of daily stock return over the contemporaneous quarter, and the standard deviation of all available quarterly analyst forecast errors from 1996–2001. We use five independent

TABLE 2
Cross-Correlation Matrix[a]

	Cost of Cap.	$\sigma(R)$	$\sigma(FE)$	*Favorable*	*Unfavorable*	*Size*	*B/M*	*Leverage*
Cost of Cap.	1.000	0.161***	0.178***	0.031**	0.006	−0.122***	0.138***	0.050***
$\sigma(R)$	0.178***	1.000	0.284***	−0.058***	0.077***	−0.069***	−0.152***	−0.184***
$\sigma(FE)$	0.122***	0.228***	1.000	−0.079***	0.094**	0.021	−0.056***	−0.089***
Favorable	−0.012	−0.044***	−0.103***	1.000	0.710***	0.061***	−0.076***	0.021
Unfavorable	−0.012	0.068***	0.042**	0.720***	1.000	0.226***	−0.167***	−0.005
Size	−0.149***	−0.127***	0.099***	0.047***	0.203***	1.000	−0.499***	0.137***
B/M	0.038***	−0.019	−0.050***	−0.180***	−0.144***	−0.242***	1.000	0.174***
Leverage	0.022	−0.167***	−0.091***	0.032**	−0.028**	0.035**	0.087***	1.000

---

[a] This table provides correlations among the variables used in subsequent tests. Spearman (Pearson) correlations are above (below) the diagonal. To be included in the sample, each firm has sufficient data to compute the variables below. The time period is from 1996 to 2001.

***, **, * Indicate two-tailed significance at the 0.01, 0.05, and 0.1 levels, respectively.

Variable Definitions:

Cost of Cap. = expected annual cost of capital calculated based on Fama-French three-factor model over the past five years' return data;
$\sigma(R)$ = standard deviation of daily stock return over the contemporaneous quarter;
$\sigma(FE)$ = standard deviation of all available quarterly analyst forecast errors from 1996 to 2001;
Favorable = favorable assessment index, constructed as mean of the content-analysis-based scaled favorable content measures for six categories of disclosures across three sources, i.e., corporations, analysts, and business press;
Unfavorable = unfavorable assessment index, constructed as mean of the content-analysis-based scaled unfavorable content measures for six categories of disclosures across three sources, i.e., corporations, analysts, and business press;
Size = natural logarithm of a firm's market capitalization of equity at the beginning of the quarter;
B/M = book equity divided by market equity at the beginning of the quarter; and
Leverage = long-term debt divided by total assets.

variables: *Favorable*, *Unfavorable*, *Size* (natural logarithm of a firm's market capitalization of equity at the beginning of the quarter), *B/M* (book equity divided by market equity at the beginning of the quarter), and *Leverage* (long-term debt divided by total asset). It is significant for all of the models at the 1 percent level. The pattern of results intuitively suggests that the regression models are analyzing a systematic effect.

Our first set of results in Table 3 consider the impact of positive and negative news disclosure by combining all three sources of disclosures, i.e., corporations, analysts, and business press, for the entire sample (Panel A) and by industry sector (Panels B–E). In interpreting the coefficients for *Favorable* and *Unfavorable* for all industries, we find highly significant effects on return volatility and analyst forecast error dispersion, but weak effect on the cost of capital. As we turn our attention to the cost of capital effects for individual industries, we find that when disclosures contain positive statements, the market anticipates this and discounts the impact of the strength of these statements. Positive news statements do not materially affect cost of capital (coefficients are weakly negative or not significant except for the pharmaceutical industry). However, when disclosures contain negative news, the market responds and the cost of capital goes up significantly in financial and technology industries, and almost significantly in the pharmaceutical industry. Return volatility and forecast dispersion effects generally hold up at the industry level of analysis.

The average cost of capital effect of unfavorable news in Panel A of Table 3 is 31 basis points, but the t-statistic is only 0.99. The reduction is the cost of capital associated with favorable news is 12 basis points with a t-statistic of -1.36. Average cost of capital effects of the magnitudes observed here are within the realm of possibility. However, when we analyze subsamples of data, the point estimates of the cost of capital effects are sometimes large, which is not surprising. Occasionally, the point estimates suggest a 200-basis-point increase or decrease, which seems unreasonably large, but these are point estimates with fair bit of standard error associated with them. Overall, the cost of capital effects appear to be modest, as expected from favorable and unfavorable news sampled across a large number of firms over a five-year period. Similar remarks apply to the estimated effects on the firms' return volatility and forecast error dispersion.

Our second set of results in Table 4, Panel A consider the impact of positive and negative news disclosure by the company's management. For both positive and negative news disclosure, information from company seems to be discounted when judged by their effect on the cost of capital. The coefficients reported are not significant for cost of capital, but the coefficients are highly significant for stock return volatility and analyst forecast errors. For positive news disclosure not reducing cost of capital, the discounting indicates a lack of credibility of management stating good news. For negative news disclosure not increasing cost of capital, the discounting suggests that bad news revealed in corporate mandatory disclosures may not be timely.

Our third set of results in Table 4, Panel B considers the impact of positive and negative news disclosure by analysts. For both positive and negative news disclosure, information from analysts appears to be discounted. Analysts on the whole are slightly less positive in their disclosures on firms, and similarly less negative (more favorable) in negative disclosure. The coefficients reported are not significant for cost of capital, but the coefficients are significant for stock return volatility and analyst forecast errors (favorable news only). Intuitively the results suggest that analysts have source credibility and/or timeliness problems. Our interpretation is that they are seen as responding to market changes after they have taken place, thus discounting their impact.

Our fourth set of results in Table 4, Panel C considers the impact of positive and negative news disclosure by the business press. Here the evidence is quite consistent. We

TABLE 3
Fama-MacBeth Regressions on the Composite News Measures[a]

Panel A: All Industries, 5,350 Observations

Independent Variables	Dependent Variable					
	Cost of Cap.		$\sigma(R)$		$\sigma(FE)$	
	Coefficient	t-stat	Coefficient	t-stat	Coefficient	t-stat
Intercept	0.1768	28.75	0.0410	19.92	1.4121	7.90
Favorable	−0.0012	−1.36	−0.0027	−10.48	−0.3319	−8.70
Unfavorable	0.0031	0.99	0.0061	5.15	0.6514	9.00
Size	−0.0047	−7.09	−0.0015	−5.08	0.0164	1.21
B/M	0.0007	0.58	−0.0030	−3.19	−0.1658	−5.09
Leverage	0.0181	1.64				
Adjusted R^2	0.0159	3.32	0.1161	6.68	0.0224	2.89

Panel B: Pharmaceutical Firms, 962 Observations

Independent Variables	Dependent Variable					
	Cost of Cap.		$\sigma(R)$		$\sigma(FE)$	
	Coefficient	t-stat	Coefficient	t-stat	Coefficient	t-stat
Intercept	0.3102	36.49	0.0724	23.41	3.6188	5.69
Favorable	−0.0064	−2.45	−0.0005	−1.10	−0.0145	−0.12
Unfavorable	0.0114	1.52	0.0022	1.52	0.3875	1.83
Size	−0.0221	−24.13	−0.0051	−16.91	−0.3809	−11.34
B/M	−0.0390	−2.96	−0.0106	−2.91	−0.6167	−2.49
Leverage	0.0724	2.57				
Adjusted R^2	0.2220	8.92	0.4668	15.38	0.3461	13.57

Panel C: Telecom Firms, 987 Observations

Independent Variables	Dependent Variable					
	Cost of Cap.		$\sigma(R)$		$\sigma(FE)$	
	Coefficient	t-stat	Coefficient	t-stat	Coefficient	t-stat
Intercept	0.2010	3.97	0.0383	11.87	2.0324	2.81
Favorable	0.0035	0.56	0.0008	1.69	−0.1732	−0.94
Unfavorable	−0.0221	−2.80	0.0001	0.08	−0.1599	−0.42
Size	−0.0059	−1.70	−0.0013	−2.64	0.2095	3.08
B/M	0.0061	1.94	−0.0021	−3.84	−0.5358	−5.85
Leverage	0.0147	0.36				
Adjusted R^2	0.1094	2.76	0.1316	4.63	−0.0885	−5.24

Panel D: Financial Firms, 3,007 Observations

Independent Variables	Dependent Variable					
	Cost of Cap.		$\sigma(R)$		$\sigma(FE)$	
	Coefficient	t-stat	Coefficient	t-stat	Coefficient	t-stat
Intercept	0.1011	9.91	0.0259	12.90	0.2671	5.46
Favorable	−0.0024	−1.80	−0.0005	−2.33	−0.0685	−3.95
Unfavorable	0.0093	3.66	0.0004	0.64	0.2944	7.15

(*continued on next page*)

TABLE 3 (continued)

Independent Variables	Dependent Variable Cost of Cap. Coefficient	 t-stat	 $\sigma(R)$ Coefficient	 t-stat	 $\sigma(FE)$ Coefficient	 t-stat
Size	0.0064	8.50	−0.0006	−1.77	−0.0113	−1.97
B/M	0.0158	1.65	0.0029	3.04	0.0674	0.89
Leverage	−0.0083	−0.62				
Adjusted R^2	0.0489	3.73	0.0618	2.34	−0.0290	−2.34

Panel E: Tech Firms, 394 Observations

Independent Variables	Dependent Variable Cost of Cap. Coefficient	 t-stat	 $\sigma(R)$ Coefficient	 t-stat	 $\sigma(FE)$ Coefficient	 t-stat
Intercept	0.1203	2.73	0.0469	7.28	6.5686	6.17
Favorable	0.0006	0.23	−0.0004	−0.41	−0.2533	−1.11
Unfavorable	0.0324	4.25	0.0073	5.29	1.4355	2.98
Size	−0.0091	−3.03	−0.0027	−5.20	−0.6057	−9.35
B/M	0.0186	0.72	−0.0150	−7.37	−3.5801	−7.35
Leverage	0.0753	2.65				
Adjusted R^2	0.1386	1.81	0.1097	3.36	0.2031	2.50

---

[a] This table provides summary statistics for the coefficients from 24 quarterly cross-sectional Fama-MacBeth regressions using the composite *Favorable* and *Unfavorable* content measures across the six categories and three sources of disclosures, i.e., corporations, analysts, and business press. Newey-West corrected time-series t-statistics are reported in the table. The time period is from 1996 to 2001.

Variable Definitions:

Cost of Cap. = expected annual cost of capital calculated based on Fama-French three-factor model over the past five years' return data;

$\sigma(R)$ = standard deviation of daily stock return over the contemporaneous quarter;

$\sigma(FE)$ = standard deviation of all available quarterly analyst forecast errors from 1996 to 2001;

Favorable = favorable assessment index, constructed as mean of the content-analysis-based scaled favorable content measures for six categories of disclosures across three sources, i.e., corporations, analysts, and business press;

Unfavorable = unfavorable assessment index, constructed as mean of the content-analysis-based scaled unfavorable content measures for six categories of disclosures across three sources, i.e., corporations, analysts, and business press;

Size = natural logarithm of a firm's market capitalization of equity at the beginning of the quarter;

B/M = book equity divided by market equity at the beginning of the quarter; and

Leverage = long-term debt divided by total assets.

find that when the business press offers both positive and negative news disclosures, there are impacts on cost of capital, return volatility, and analyst forecast error dispersion. Positive news disclosure decreases cost of capital, negative news disclosure increases it. The return volatility and forecast dispersion effects are also highly significant in the predicted direction. Intuitively, the strength of this finding suggests that the credibility of news disclosure by the press is higher than that for companies or for analysts, and in the case of analysts, press disclosure is seen as more timely. The press is seen as less affected by incentive concerns (agency), and this is reflected in the consistency of evidence.

The lack of effect on the firm's cost of capital using disclosures from corporations, and the findings of the effect on the firm's cost of capital using disclosures from the business press are also consistent with evidence presented in Li (2008), who concludes that the

TABLE 4
Fama-MacBeth Regressions by the Source of News[a]

Panel A: Disclosure Source is Corporate Disclosures

	Dependent Variable					
	Cost of Cap.		***$\sigma(R)$***		***$\sigma(FE)$***	
Independent Variables	**Coefficient**	**t-stat**	**Coefficient**	**t-stat**	**Coefficient**	**t-stat**
Intercept	0.1825	26.94	0.0422	19.89	1.6587	6.62
Favorable	0.0003	0.39	−0.0015	−6.08	−0.2123	−6.57
Unfavorable	−0.0036	−1.59	0.0025	3.06	0.3429	9.44
Size	−0.0047	−7.17	−0.0014	−4.09	0.0010	0.06
B/M	0.0017	0.60	−0.0033	−3.31	−0.2583	−4.21
Leverage	0.0113	1.17				
Adjusted R^2	0.0161	3.13	0.1008	5.44	0.0391	3.46

Panel B: Disclosure Source is Analyst Reports from Investext

	Dependent Variable					
	Cost of Cap.		***$\sigma(R)$***		***$\sigma(FE)$***	
Independent Variables	**Coefficient**	**t-stat**	**Coefficient**	**t-stat**	**Coefficient**	**t-stat**
Intercept	0.1771	30.98	0.0398	21.99	0.9812	7.81
Favorable	0.0000	0.00	−0.0017	−7.98	−0.0758	−2.17
Unfavorable	0.0022	0.27	0.0051	6.07	0.0175	0.17
Size	−0.0051	−9.15	−0.0015	−4.23	0.0548	6.30
B/M	0.0003	0.22	−0.0029	−3.31	−0.1128	−3.09
Leverage	0.0159	1.60				
Adjusted R^2	0.0234	3.54	0.0965	5.19	−0.0029	−0.71

Panel C: Disclosure Source is Business Press

	Dependent Variable					
	Cost of Cap.		***$\sigma(R)$***		***$\sigma(FE)$***	
Independent Variables	**Coefficient**	**t-stat**	**Coefficient**	**t-stat**	**Coefficient**	**t-stat**
Intercept	0.1757	36.42	0.0367	19.5	0.5192	6.82
Favorable	−0.0013	−2.82	−0.0004	−2.80	−0.0823	−3.52
Unfavorable	0.0026	1.85	0.0017	6.07	0.2635	8.38
Size	−0.0045	−5.53	−0.0014	−4.26	0.0448	2.49
B/M	0.0011	0.77	−0.0026	−3.09	−0.0766	−2.69
Leverage	0.0167	1.46				
Adjusted R^2	0.0166	4.25	0.0779	3.92	−0.0018	−0.41

---

[a] This table provides summary statistics for the coefficients from 24 quarterly cross-sectional Fama-MacBeth regressions using the *Favorable* and *Unfavorable* content measures for each source of disclosure, i.e., corporations, analysts, and business press. The content measures are summed across the six categories of disclosures. Newey-West corrected time-series t-statistics are reported in the table. The time period is from 1996 to 2001.

Variable Definitions:

Cost of Cap. = expected annual cost of capital calculated based on Fama-French three-factor model over the past five years' return data;

$\sigma(R)$ = standard deviation of daily stock return over the contemporaneous quarter;

(*continued on next page*)

TABLE 4 (continued)

$\sigma(FE)$ = standard deviation of all available quarterly analyst forecast errors from 1996 to 2001;
Favorable = favorable assessment index, constructed as mean of the content-analysis-based scaled favorable content measures for six categories of disclosures across three sources, i.e., corporations, analysts, and business press;
Unfavorable = unfavorable assessment index, constructed as mean of the content-analysis-based scaled unfavorable content measures for six categories of disclosures across three sources, i.e., corporations, analysts, and business press;
Size = natural logarithm of a firm's market capitalization of equity at the beginning of the quarter;
B/M = book equity divided by market equity at the beginning of the quarter; and
Leverage = long-term debt divided by total assets.

General Inquirer does not accurately analyze the tone of MD&As in corporate filings. Conversely, the GI accurately analyzes documents from the general news media and business press. Li's (2008) findings suggest that the measures derived from content analysis for corporate disclosures are more likely to be noisy than those from the business press. However, our result showing an effect of the content in corporate disclosures and business press on return volatility and analyst forecast dispersion suggests that the informativeness of corporate disclosures is not entirely lost due to the noise and bias in disclosures stemming from management incentives.

We next examine whether the effect of disclosures by the management, analysts, and business press have differential impact on the firms' cost of capital, volatility, and analyst forecast error dispersion when it relates to small versus large market capitalization stocks. Table 5 reports separate analysis for small and large firms. To designate a firm as small or large, we sort the sample firms by their equity market capitalization. One-half of the sample is defined as small firms and the other half is labeled large firms.

Findings for small firms are based on 2,682 firm-year observations. The analysis is portrayed in three panels. Panel A examines the relation between management disclosures and the cost of capital as well as return volatility and forecast error variability. Panels B and C contain analysis similar to that in Panel A except that the source of disclosures is analyst reports in Panel B and business press in Panel C.

Results in Table 5 show that regardless of the source, favorable disclosures reduce the cost of capital, whereas unfavorable disclosures are associated with higher cost of capital. For small firms, information is relatively scarce, and disclosures, favorable or unfavorable, enable the market participants to learn more about the risks facing the firm, which are reflected in the firm's cost of capital. The results lead us to similar conclusions with respect to return volatility and analyst forecast errors.

Results in Panel A suggest, somewhat paradoxically, that the market deems all three sources of information as credible enough to influence the expected cost of equity capital of the firms experiencing favorable and unfavorable coverage. The result is due likely to the fact that small firms have considerable amount of uncertainty swirling around their prospects. Therefore, new disclosures influence investors' assessment regardless of the source, which results in disclosures being associated with the three measures of risk.

In Panels D, E, and F of Table 5, results for large cap stocks reveal a different picture for cost of capital effects, but the other two measures of risk, volatility and analyst forecast dispersion, show results consistent with the hypothesis. Here, cost of capital is less systematically associated with disclosure content. In the presence of considerable amount of information already available for large cap stocks, new disclosures are not particularly helpful in updating investor beliefs. As a result, there is less consistent indication of an impact of disclosures on large firms' cost of capital. In fact, in the case of management disclosures,

TABLE 5
Fama-MacBeth Regressions by the Source of News: Small versus Large[a]

Panel A: Disclosure Source is Corporate Disclosures: Small Firms, 2,682 Observations

	Dependent Variable					
Independent Variables	*Cost of Cap.*		$\sigma(R)$		$\sigma(FE)$	
	Coefficient	**t-stat**	**Coefficient**	**t-stat**	**Coefficient**	**t-stat**
Intercept	0.1579	11.99	0.0581	12.90	0.4090	1.74
Favorable	−0.0048	−4.51	−0.0011	−1.85	−0.2327	−3.34
Unfavorable	0.0063	2.93	0.0010	0.65	0.6716	2.87
Size	0.0054	3.28	−0.0045	−5.06	0.1086	3.15
B/M	−0.0098	−3.03	−0.0050	−2.55	−0.2893	−3.75
Leverage	0.0090	0.52				
Adjusted R^2	0.0197	1.31	0.1350	5.06	0.0676	1.59

Panel B: Disclosure Source is Analyst Reports from Investext: Small Firms, 2,682 Observations

	Dependent Variable					
Independent Variables	*Cost of Cap.*		$\sigma(R)$		$\sigma(FE)$	
	Coefficient	**t-stat**	**Coefficient**	**t-stat**	**Coefficient**	**t-stat**
Intercept	0.1410	10.64	0.0550	11.47	0.2036	1.94
Favorable	−0.0040	−2.37	−0.0016	−8.75	−0.0376	−2.56
Unfavorable	0.0169	2.68	0.0055	6.32	0.1480	2.85
Size	0.0047	2.92	−0.0045	−6.14	0.1121	4.08
B/M	−0.0100	−2.79	−0.0056	−2.30	−0.2716	−2.38
Leverage	0.0069	0.47				
Adjusted R^2	0.0220	1.83	0.1423	5.35	−0.0194	−3.68

Panel C: Disclosure Source is Business Press: Small Firms, 2,682 Observations

	Dependent Variable					
Independent Variables	*Cost of Cap.*		$\sigma(R)$		$\sigma(FE)$	
	Coefficient	**t-stat**	**Coefficient**	**t-stat**	**Coefficient**	**t-stat**
Intercept	0.1395	11.69	0.0533	10.49	0.2948	1.83
Favorable	−0.0029	−2.57	−0.0002	−0.71	−0.0771	−3.21
Unfavorable	0.0093	3.33	0.0018	4.56	0.3783	7.01
Size	0.0048	2.40	−0.0048	−5.69	0.0629	1.97
B/M	−0.0091	−2.76	−0.0052	−2.30	−0.2673	−2.16
Leverage	0.0084	0.51				
Adjusted R^2	0.0204	1.82	0.1263	4.10	0.0736	3.49

Panel D: Disclosure Source is Corporate Disclosures: Large Firms, 2,668 Observations

	Dependent Variable					
Independent Variables	*Cost of Cap.*		$\sigma(R)$		$\sigma(FE)$	
	Coefficient	**t-stat**	**Coefficient**	**t-stat**	**Coefficient**	**t-stat**
Intercept	0.1938	12.92	0.0316	22.94	4.4315	15.02
Favorable	0.0039	4.02	−0.0013	−11.77	−0.2558	−10.90

(*continued on next page*)

TABLE 5 (continued)

	Dependent Variable					
Independent Variables	*Cost of Cap.* **Coefficient**	**t-stat**	*σ(R)* **Coefficient**	**t-stat**	*σ(FE)* **Coefficient**	**t-stat**
Unfavorable	−0.0083	−2.86	0.0024	5.10	0.2766	7.31
Size	−0.0077	−7.33	−0.0003	−1.56	−0.2357	−12.55
B/M	0.0135	2.35	−0.0024	−3.46	−0.3444	−7.02
Leverage	−0.0041	−0.41				
Adjusted R^2	0.0554	4.14	0.0659	6.15	0.0192	3.09

Panel E: Disclosure Source is Analyst Reports from Investext: Large Firms, 2,668 Observations

	Dependent Variable					
Independent Variables	*Cost of Cap.* **Coefficient**	**t-stat**	*σ(R)* **Coefficient**	**t-stat**	*σ(FE)* **Coefficient**	**t-stat**
Intercept	0.2060	21.41	0.0290	18.16	4.4686	12.77
Favorable	0.0017	1.68	−0.0009	−2.80	−0.2223	−6.78
Unfavorable	−0.0038	−0.68	0.0028	4.54	0.1372	2.09
Size	−0.0087	−10.77	−0.0003	−1.48	−0.2357	−9.58
B/M	0.0080	3.98	−0.0017	−3.61	−0.2793	−5.47
Leverage	−0.0006	−0.07				
Adjusted R^2	0.0422	3.42	0.0386	7.59	0.0007	0.07

Panel F: Disclosure Source is Business Press: Large Firms, 2,668 Observations

	Dependent Variable					
Independent Variables	*Cost of Cap.* **Coefficient**	**t-stat**	*σ(R)* **Coefficient**	**t-stat**	*σ(FE)* **Coefficient**	**t-stat**
Intercept	0.2256	32.11	0.0245	20.22	3.4893	9.69
Favorable	−0.0006	−0.53	−0.0003	−2.48	−0.1671	−3.91
Unfavorable	−0.0030	−1.59	0.0016	7.25	0.1983	4.76
Size	−0.0095	−20.33	−0.0001	−0.34	−0.2236	−10.05
B/M	0.0077	3.48	−0.0014	−2.60	−0.1218	−3.58
Leverage	−0.0056	−0.58				
Adjusted R^2	0.0386	3.03	0.0391	4.96	−0.0241	−6.38

---

[a] This table provides summary statistics for the coefficients from 24 quarterly cross-sectional Fama-MacBeth regressions using the *Favorable* and *Unfavorable* content measures for each source of disclosure, i.e., corporations, analysts, and business press. Observations are sorted into two groups by firms' market capitalization. The content measures are summed across the six categories of disclosures. Newey-West corrected time-series t-statistics are reported in the table. The time period is from 1996 to 2001.

Variable Definitions:

Cost of Cap. = expected annual cost of capital calculated based on Fama-French three-factor model over the past five years' return data;

σ(R) = standard deviation of daily stock return over the contemporaneous quarter;

σ(FE) = standard deviation of all available quarterly analyst forecast errors from 1996 to 2001;

Favorable = favorable assessment index, constructed as mean of the content-analysis-based scaled favorable content measures for six categories of disclosures across three sources, i.e., corporations, analysts, and business press;

(*continued on next page*)

TABLE 5 (continued)

Unfavorable = unfavorable assessment index, constructed as mean of the content-analysis-based scaled unfavorable content measures for six categories of disclosures across three sources, i.e., corporations, analysts, and business press;
Size = natural logarithm of a firm's market capitalization of equity at the beginning of the quarter;
B/M = book equity divided by market equity at the beginning of the quarter; and
Leverage = long-term debt divided by total assets.

---

favorable disclosures are associated with significantly higher, not lower, cost of capital. However, favorable assessment in analyst reports and business press is not associated with a significant change in the cost of capital. Similarly, unfavorable disclosures are not consistently associated with higher cost of capital. Overall, the cost of capital effect of disclosures is not observed consistently for large stocks. In contrast to the cost of capital effect, disclosures' effect on return volatility and analyst forecast error dispersion continues to be significant and as predicted for the large firms. Thus, only the cost of capital effect is puzzling.

Table 6 reports results of analyzing the impact of disclosures by source the cost of capital of firms with low and high institutional ownership. Like with size of the firm, we define high and low institutional ownership as firms in the top and bottom half when the sample firms are ranked on institutional ownership.

The results are similar to those we obtain for small and large cap stocks. This is not surprising because one expects a positive correspondence between institutional ownership and firm size and we just discussed findings for firms partitioned on the basis of firm size.

V. IMPLICATIONS FOR MANAGEMENT AND FUTURE RESEARCH

Regulation of disclosure and the materiality of information that a reasonable investor would want to consider in making an investment decision are subjects of intense scrutiny. The materiality of information relates to the importance and impact of corporate information and disclosures having on the earnings of the company, and to how disclosure should reach the market and investors (Brown 1998). As the disclosure system has grown in size with the importance of evaluation of disclosure by other information intermediaries, analysts, business press, and the investors themselves, information asymmetries and agency conflicts between participants creates greater demand for reporting and assurance of disclosure.

The research reported here is the first to document systematic evidence of the effect of disclosure on various risk measures, including the cost of capital. Our analysis of disclosure is based on a virtually exhaustive set of sources from the print medium, segregated by disclosure source—corporate, analysts, and news stories published in the business press. We find support that the market weighs disclosures according to the credibility of source, and that, on average, the market filters out and deemphasizes noisy disclosures by all participants. Our results find that negative news disclosure is strongly weighted by the market, and positive news is discounted as firms and investment analysts have incentives to skew disclosure. Our results suggest that corporations' and analysts' communications with the investment community are not credible. Specifically, analysts' reports are heavily discounted, suggesting that either they are not credible or that they are seen as responding to market changes after they have taken place, discounting their impact.

Problems of agency and the incentives for the business press to skew disclosure are more diffuse for journalists than for companies or analysts, and our results reflect this more

TABLE 6
Fama-MacBeth Regressions by the Source of News: Low Institutional Ownership versus High Institutional Ownership[a]

Panel A: Disclosure Source is Corporate Disclosures: Low Institutional Ownership Firms, 2,501 Observations

	Dependent Variable					
Independent Variables	*Cost of Cap.*		$\sigma(R)$		$\sigma(FE)$	
	Coefficient	**t-stat**	**Coefficient**	**t-stat**	**Coefficient**	**t-stat**
Intercept	0.1741	15.60	0.0390	23.26	1.2304	4.20
Favorable	−0.0020	−0.86	−0.0014	−3.41	−0.1754	−2.16
Unfavorable	0.0011	0.21	0.0029	2.74	0.1611	0.91
Size	−0.0040	−2.41	−0.0015	−4.88	0.0918	3.28
B/M	−0.0023	−1.31	−0.0009	−1.64	−0.2536	−4.56
Leverage	0.0816	3.16				
Adjusted R^2	0.0286	2.13	0.0651	6.97	0.1010	2.97

Panel B: Disclosure Source is Analyst Reports from Investext: Low Institutional Ownership Firms, 2,501 Observations

	Dependent Variable					
Independent Variables	*Cost of Cap.*		$\sigma(R)$		$\sigma(FE)$	
	Coefficient	**t-stat**	**Coefficient**	**t-stat**	**Coefficient**	**t-stat**
Intercept	0.1613	20.12	0.0382	28.35	−0.2043	−1.56
Favorable	−0.0038	−2.53	−0.0013	−3.50	−0.1192	−7.06
Unfavorable	0.0145	2.77	0.0044	3.15	0.5022	8.82
Size	−0.0035	−3.40	−0.0017	−5.70	0.1554	9.18
B/M	0.0014	0.76	−0.0007	−1.48	−0.0452	−1.52
Leverage	0.0726	3.55				
Adjusted R^2	0.0338	1.94	0.0803	9.45	0.0475	3.58

Panel C: Disclosure Source is Business Press: Low Institutional Ownership Firms, 2,501 Observations

	Dependent Variable					
Independent Variables	*Cost of Cap.*		$\sigma(R)$		$\sigma(FE)$	
	Coefficient	**t-stat**	**Coefficient**	**t-stat**	**Coefficient**	**t-stat**
Intercept	0.1550	24.64	0.0353	37.71	−0.2364	−2.95
Favorable	−0.0026	−3.41	0.0000	−0.13	−0.0034	−0.12
Unfavorable	0.0078	2.86	0.0013	4.88	0.2331	5.18
Size	−0.0029	−2.12	−0.0016	−5.40	0.1379	9.48
B/M	0.0015	0.82	−0.0006	−1.12	−0.0407	−1.53
Leverage	0.0710	3.42				
Adjusted R^2	0.0272	5.07	0.0565	3.88	0.0588	3.43

(*continued on next page*)

TABLE 6 (continued)

Panel D: Disclosure Source is Corporate Disclosures: High Institutional Ownership Firms, 2,489 Observations

Independent Variables	Dependent Variable *Cost of Cap.* Coefficient	t-stat	*σ(R)* Coefficient	t-stat	*σ(FE)* Coefficient	t-stat
Intercept	0.1884	15.18	0.0377	11.88	2.7244	7.95
Favorable	0.0037	3.88	−0.0015	−9.39	−0.2106	−8.82
Unfavorable	−0.0101	−4.87	0.0023	4.07	0.3652	7.58
Size	−0.0064	−5.35	−0.0007	−1.83	−0.1149	−4.35
B/M	0.0176	2.38	−0.0082	−4.28	−0.3886	−1.61
Leverage	−0.0292	−2.85				
Adjusted R^2	0.0558	2.97	0.1406	5.44	−0.0002	−0.04

Panel E: Disclosure Source is Analyst Reports from Investext: High Institutional Ownership Firms, 2,489 Observations

Independent Variables	Dependent Variable *Cost of Cap.* Coefficient	t-stat	*σ(R)* Coefficient	t-stat	*σ(FE)* Coefficient	t-stat
Intercept	0.1860	10.89	0.0352	12.40	3.6303	9.08
Favorable	0.0043	3.49	−0.0011	−5.70	−0.1911	−5.55
Unfavorable	−0.0071	−1.66	0.0039	6.21	−0.0670	−0.85
Size	−0.0074	−4.31	−0.0008	−1.93	−0.1295	−3.62
B/M	0.0117	2.35	−0.0093	−4.03	−0.6093	−1.81
Leverage	−0.0233	−2.49				
Adjusted R^2	0.0601	2.84	0.1054	4.55	0.0051	0.43

Panel F: Disclosure Source is Business Press: High Institutional Ownership Firms, 2,489 Observations

Independent Variables	Dependent Variable *Cost of Cap.* Coefficient	t-stat	*σ(R)* Coefficient	t-stat	*σ(FE)* Coefficient	t-stat
Intercept	0.2064	16.88	0.0329	9.55	2.4989	5.34
Favorable	0.0018	1.47	−0.0004	−3.34	−0.2040	−3.95
Unfavorable	−0.0054	−2.30	0.0017	6.68	0.3165	6.60
Size	−0.0077	−4.94	−0.0008	−1.78	−0.1249	−5.60
B/M	0.0126	2.39	−0.0087	−3.85	−0.5894	−2.03
Leverage	−0.0232	−2.32				
Adjusted R^2	0.0529	2.46	0.1092	6.11	−0.0253	−3.53

---

[a] This table provides summary statistics for the coefficients from quarterly cross-sectional Fama-MacBeth regressions using the *Favorable* and *Unfavorable* content measures for each source of disclosure, i.e., corporations, analysts, and business press. Observations are sorted into two groups by firms' institutional ownership. The content measures are summed across the six categories of disclosures. Newey-West corrected time-series t-statistics are reported in the table. The time period is from 1996 to 2001.

Variable Definitions:

Cost of Cap. = expected annual cost of capital calculated based on Fama-French three-factor model over the past five years' return data;

(*continued on next page*)

TABLE 6 (continued)

$\sigma(R)$ = standard deviation of daily stock return over the contemporaneous quarter;
$\sigma(FE)$ = standard deviation of all available quarterly analyst forecast errors from 1996 to 2001;
Favorable = favorable assessment index, constructed as mean of the content-analysis-based scaled favorable content measures for six categories of disclosures across three sources, i.e., corporations, analysts, and business press;
Unfavorable = unfavorable assessment index, constructed as mean of the content-analysis-based scaled unfavorable content measures for six categories of disclosures across three sources, i.e., corporations, analysts, and business press;
Size = natural logarithm of a firm's market capitalization of equity at the beginning of the quarter;
B/M = book equity divided by market equity at the beginning of the quarter; and
Leverage = long-term debt divided by total assets.

independent intermediary role—positive news disclosure decreases cost of capital, and negative news disclosure increases it. Intuitively this finding suggests the credibility of news disclosure by the press is higher than that for companies or analysts.

The strength and pattern of our results suggest that there are important economic reasons for companies to understand more precisely how disclosure affects the firm's cost of capital and other measures of risk. Pursuant to securities regulation and business norms, companies have a duty to disclose corporate information and developments through timely and accurate disclosure. That we have found evidence showing that the market discounts disclosure by source credibility suggests that companies should give high priority to developing appropriate and complete communications policies. Skewing disclosure of critical corporate information or development has economic risks. Asking whether consistent, accurate, and complete disclosure is best accomplished through centralizing the corporate communication process, or decentralizing it with strong behavioral norms and safeguards, is a key senior-management question. An implication of our results is that companies cannot depend on analysts to enhance source credibility or timeliness of disclosure. Analysts have their own credibility and timeliness problems, and are subjects of increased scrutiny in disclosure reform.

We conclude by noting that financial reporting and disclosure will continue to be an important field of empirical inquiry. There are numerous factors with the potential to alter disclosure practices and financial reporting—technological innovation, changes in the business economics of audit firms and analysts, globalization of capital markets, and changes in disclosure channels and the number and type of information intermediaries—that continue to reshape disclosure and financial reporting practices. They create new and exciting opportunities for research.

REFERENCES

Amihud, Y., and H. Mendelson. 1986. Asset pricing and the bid-ask spread. *Journal of Financial Economics* 17: 223–249.

Ball, R., and S. Kothari. 1989. Nonstationary expected returns: Implications for tests of market efficiency and serial correlation in returns. *Journal of Financial Economics* 25: 51–74.

Barry, C., and S. Brown. 1984. Differential information and the small firm effect. *Journal of Financial Economics* 13: 283–294.

Beaver, W. 1998. *Financial Reporting: An Accounting Revolution.* 3rd edition. Upper Saddle River, NJ: Prentice Hall.

Botosan, C. 1997. Disclosure level and the cost of capital. *The Accounting Review* 72: 323–350.

———, and M. Plumlee. 2002. A re-examination of disclosure level and the expected cost of equity capital. *Journal of Accounting Research* 40: 21–41.

Bradshaw, M., S. Richardson, and R. Sloan. 2003. Pump and dump: An empirical analysis of the relation between corporate financing activities and sell-side analyst research. Working paper, University of Michigan.

Brown, R., Jr. 1998. *The Regulation of Corporate Disclosure*. 3rd edition. New York, NY: Aspen Publishers, Inc.

Bushee, B., J. Core, W. Guay, and S. Hamm. 2008. The role of business press as an information intermediary. Working paper, The Wharton School, University of Pennsylvania.

Bushman, R., and A. Smith. 2001. Financial accounting research and corporate governance. *Journal of Accounting and Economics* 31: 237–333.

Chan, K. 1988. On the contrarian investment strategy. *Journal of Business* 61: 147–163.

Chordia, T., R. Roll, and A. Subrahmanyam. 2001. Market liquidity and trading activity. *The Journal of Finance* 56: 501–530.

Christie, A. 1982. The stochastic behavior of common stock variances: Value, leverage and interest rate effects. *Journal of Financial Economics* 10: 407–432.

Clarke, J., A. Khorana, A. Patel, and P. Rau. 2007. The impact of all-star analyst job changes on their coverage choices and investment banking deal flow. *Journal of Financial Economics* 84: 713–737.

Claus, J., and J. Thomas. 2001. Equity premia as low as three percent? Evidence from analysts' earnings forecasts for domestic and international stock markets. *The Journal of Finance* 56: 1629–1666.

Core, J. 2001. A review of the empirical disclosure literature. *Journal of Accounting and Economics* 31: 441–456 discussion.

———, W. Guay, and R. Verdi. 2008. Is accruals quality a priced factor? *Journal of Accounting and Economics* 46: 2–22.

Dechow, P., A. Hutton, and R. Sloan. 2000. The relation between analysts' long-term earnings forecasts and stock price performance following equity offerings. *Contemporary Accounting Research* 17: 1–32.

Diamond, D., and R. Verrecchia. 1991. Disclosure, liquidity, and the cost of capital. *The Journal of Finance* 46: 1325–1355.

Easley, D., S. Hvidkjaer, and M. O'Hara. 2002. Is information risk a determinant of asset returns? *The Journal of Finance* 57: 2185–2222.

———, and M. O'Hara. 2004. Information and the cost of capital. *The Journal of Finance* 59: 1553–1583.

Easton, P. 2004. PE ratios, PEG ratios, and estimating the implied expected rate of return on equity capital. *The Accounting Review* 79: 73–95.

Fama, E., and K. French. 1993. Common risk factors in the returns on stocks and bonds. *Journal of Financial Economics* 33: 3–56.

Fields, T., T. Lys, and L. Vincent. 2001. Empirical research on accounting choice. *Journal of Accounting and Economics* 31: 255–307.

Francis, J., R. LaFond, P. Olsson, and K. Schipper. 2005. The market pricing of accruals quality. *Journal of Accounting and Economics* 39: 295–327.

Frankel, R., S. Kothari, and J. Weber. 2006. Determinants of the informativeness of analyst research. *Journal of Accounting and Economics* 41: 29–54.

French, K., G. Schwert, and R. Stambaugh. 1987. Expected stock returns and volatility. *Journal of Financial Economics* 19: 3–29.

Galai, D., and R. Masulis. 1976. The option pricing model and the risk factor of stock. *Journal of Financial Economics* 3: 53–81.

Gebhardt, W., C. Lee, and B. Swaminathan. 2001. Toward an implied cost of capital. *Journal of Accounting Research* 39: 135–176.

Glosten, L., and P. Milgrom. 1985. Bid, ask, and transaction prices in a specialist market with heterogeneously informed traders. *Journal of Financial Economics* 14: 71–100.

Graham, J., C. Harvey, and S. Rajgopal. 2005. The economic implications of corporate financial reporting. *Journal of Accounting and Economics* 40: 3–73.

Guay, W., S. Kothari, and S. Shu. 2006. Properties of implied cost of capital using analysts' forecasts. Working paper, MIT Sloan School of Management.

Hamburger, T., and J. Burns. 2002. Senate panel grills analysts who pushed Enron stock even as the firm faltered. *World Street Journal* (February 28): A3.

Healy, P., and K. Palepu. 2001. Information asymmetry, corporate disclosure, and the capital markets: A review of the empirical literature. *Journal of Accounting and Economics* 31: 405–441.

Kasznik, R., and B. Lev. 1995. To warn or not to warn: Management disclosures in the face of an earnings surprise. *The Accounting Review* 70: 113–134.

Kellogg, R. 1984. Accounting activities, security prices, and class action lawsuits. *Journal of Accounting and Economics* 6: 185–204.

Klein, R., and V. Bawa. 1976. The effect of estimation risk on optimal portfolio choice. *Journal of Financial Economics* 3: 215–231.

Kolasinski, A., and S. Kothari. 2008. Investment banking and analyst objectivity: Evidence from analysts affiliated with M&A advisors. *Journal of Financial and Quantitative Analysis* 43: 817–842.

Kothari, S., S. Shu, and P. Wysocki. 2009. Do managers withhold bad news? *Journal of Accounting Research* 47: 241–276.

Lambert, R., C. Leuz, and R. Verrecchia. 2007. Accounting information, disclosure, and the cost of capital. *Journal of Accounting Research* 45: 385–420.

Lang, M., and R. Lundholm. 1996. Corporate disclosure policy and analyst behavior. *The Accounting Review* 71: 467–492.

Leuz, C., and P. Wysocki. 2008. Economic consequences of financial reporting and disclosure regulation: A review and suggestions for future research. Working paper, MIT Sloan School of Management.

Li, F. 2008. The determinants and information contents of the forward-looking statements in corporate filings—A naïve Bayesian machine learning approach. Working paper, University of Michigan.

Lin, H., and M. McNichols. 1998. Underwriting relationships, analysts' earnings forecasts and investment recommendations. *Journal of Accounting and Economics* 25: 101–127.

Liu, M., and P. Wysocki. 2007. Cross-sectional determinants of information quality proxies and cost of capital measures. Working paper, MIT Sloan School of Management.

Ljungqvist, A., F. Marston, L. Starks, K. Wei, and H. Yan. 2007. Conflicts of interest in sell-side research and the moderating role of institutional investors. *Journal of Financial Economics* 85: 420–456.

Lys, T., and S. Sohn. 1990. The association between revisions of financial analysts' earnings forecasts and security price changes. *Journal of Accounting and Economics* 13: 341–363.

Michaely, R., and K. Womack. 1999. Conflict of interest and the credibility of underwriter analyst recommendations. *Review of Financial Studies* 12: 573–608.

Ng, J., R. Verrecchia, and J. Weber. 2009. Firm performance measures and adverse selection. Working paper, MIT Sloan School of Management.

O'Brien, P., M. McNichols, and H. Lin. 2005. Analyst impartiality and investment banking relationships. *Journal of Accounting Research* 43: 623–650.

Ohlson, J., and B. Juettner-Nauroth. 2005. Expected EPS and EPS growth as determinants of value. *Review of Accounting Studies* 10: 349–365.

Pastor, L., and R. Stambaugh. 2003. Liquidity risk and expected stock returns. *The Journal of Political Economy* 111: 642–685.

Riloff, E. 1993. Automatically constructing a dictionary for information extraction tasks. In *Proceedings of the Eleventh National Conference on Artificial Intelligence*, 811–816, AAAI Press/The MIT Press.

———, and L. Hollaar. 1996. Text databases and information retrieval. *ACM Computing Surveys* 28: 133–135.

Semin, G. R., and K. Fiedler, eds. 1992. *Language, Interaction and Social Cognition.* Thousand Oaks, CA: Sage Publications.

———, and ———, eds. 1996. *Applied Social Psychology.* Thousand Oaks, CA: Sage Publications.

Skinner, D. 1994. Why firms voluntarily disclose bad news. *Journal of Accounting Research* 32: 38–60.

———. 1997. Earnings disclosures and stockholder lawsuits. *Journal of Accounting and Economics* 23: 249–283.

Verrecchia, R. 2001. Essays on disclosure. *Journal of Accounting and Economics* 32: 97–180.

Watts, R., and J. Zimmerman. 1986. *Positive Accounting Theory*. Englewood Cliffs, NJ: Prentice Hall.

THE ACCOUNTING REVIEW
Vol. 84, No. 5
2009
pp. 1671–1694

American Accounting Association
DOI: 10.2308/accr.2009.84.5.1671

Reciprocity and the Effectiveness of Optimal Agency Contracts

Xi (Jason) Kuang
Georgia Institute of Technology

Donald V. Moser
University of Pittsburgh

ABSTRACT: Optimal agency contracts pay the lowest wage necessary to induce profit-maximizing effort. Employees could view such contracts as violating reciprocity because, relative to more reciprocal contracts, they offer a lower wage in exchange for higher effort. Consequently, the profit-maximizing effectiveness of optimal contracts could be impaired if employees reject them or reduce their effort. We use experimental labor markets to examine (1) how employees respond to an optimal versus a suboptimal reciprocity-based contract when each contract is the only contract available, (2) how employees respond to these contracts when firms choose which one to offer, (3) whether the firms' contract offers depend on employees' reactions to those offers, and (4) how employees and firms react to a hybrid contract that incorporates features of both contracts. We find that the optimal contract is less effective than agency analysis predicts, the reciprocity-based contract can be equally effective, and the hybrid contract dominates a market in which all three contracts are available. Implications of these results are discussed.

Keywords: *reciprocity; gift exchange; agency theory; optimal contract.*

Data Availability: *Please contact the authors.*

I. INTRODUCTION

Agency theory is one of the most important theoretical paradigms in management accounting research (Indjejikian 1999; Lambert 2001). The main focus of agency research is on deriving optimal incentive contracts in various control environments

---

We appreciate the helpful comments of Jason Brown, Alan Douglas, Harry Evans, Vicky Hoffman, Kathryn Kadous, Karen Shastri, Kristy Towry, Greg Waymire, and workshop participants at Emory University, University of Pittsburgh, and University of Waterloo on this version of the paper. We also gratefully acknowledge helpful comments from Jake Birnberg, Bryan Church, Harry Evans, Mei Feng, Lynn Hannan, Vicky Hoffman, Karim Jamal, Kathryn Kadous, Grace Pownall, Arnie Schneider, Tom Scott, Dhinu Srinivasan, Roberto Weber, and workshop participants at Cornell University, Emory University, Florida State University, Georgia Institute of Technology, University of Alberta, University of Pittsburgh, University of Southern California, The University of Texas, and the 2005 ABO Conference on an earlier version of the paper that included a subset of the data included in the paper.

Editor's note: Accepted by Steven Kachelmeier, with thanks to Mark Nelson for serving as editor on a previous version.

Submitted: November 2005
Accepted: November 2008
Published Online: September 2009

(Baiman 1982). The goal of such contracts is to maximize firm profit, taking into account employees' rationality and incentive compatibility constraints. Optimal contracts achieve this goal by offering employees the lowest wage necessary to induce the effort that maximizes firm profits.

However, in addition to the financial incentives in employment contracts, prior experimental research in accounting (Hannan 2005) and economics (Fehr et al. 1993; Fehr et al. 1998; Hannan et al. 2002) has documented a reciprocity norm between firms and employees. Such reciprocity between firms and employees has been labeled "gift exchange" because, inconsistent with the predictions of conventional economic analysis, firms give employees a gift of a wage above the market-clearing level, and employees reciprocate with a gift of effort above the enforceable level (Akerlof 1982). This research has also shown that, on average, this type of reciprocity yields higher firm profit than if, as economic theory prescribes, firms pay only the market-clearing wage and employees provide only the enforceable effort level.

Preferences for reciprocity of the type documented in gift-exchange studies play no role in standard agency analysis. However, such preferences could reduce the profit-maximizing effectiveness of optimal agency contracts if employees expect the firm to maintain a reciprocal relationship with them and retaliate when the firm fails to do so (Robinson et al. 1994; Rousseau 1995). Because optimal agency contracts are designed to pay the lowest wage necessary to induce the effort that maximizes firm profit, they could be perceived as violating the reciprocity norm. If employees punish firms that offer theoretically optimal contracts for violating reciprocity by rejecting such contracts or by accepting them and reducing effort, then the profit-maximizing effectiveness of the contracts will be impaired.

We conduct three experiments using labor markets to address a number of related research questions: (1) how employees respond to being offered an optimal agency contract versus a theoretically suboptimal reciprocity-based gift-exchange contract when each contract is the only contract available in the market, (2) how employees respond to these two contracts in a market in which the firm chooses which one to offer, (3) whether firms' contract offers depend on how employees react to those offers, and (4) how employees and firms react to a hybrid contract that incorporates important features of both the theoretically optimal and reciprocity-based contracts.

The results of our first experiment provide evidence that the optimal agency contract for our experimental setting is somewhat effective as a profit-maximizing contract in that, consistent with the agency theory prediction, it yields significantly higher firm profit than the suboptimal reciprocity-based gift-exchange contract when each contract is the only contract available in the market. However, some employees react negatively to the optimal contract and provide less than the optimal effort level even though it is costly for them to do so. The results of our second experiment show that when both contracts are available and firms choose which one to offer, firm profit from the optimal contract decreases significantly because employees punish firms for offering it, either by rejecting the contract entirely or by accepting it and reducing effort. Consequently, in later periods of the experiment, firm profit under the reciprocity-based gift-exchange contract is statistically equivalent to that under the optimal contract. Moreover, in later periods, firms begin offering the reciprocity-based contract more often than the optimal contract. These results are consistent with our expectation that the profit-maximizing effectiveness of optimal agency contracts can be impaired relative to the standard agency prediction if they are viewed as violating the reciprocity norm.

The results of our third experiment show that a new hybrid contract that combines the "forcing" feature of the optimal agency contract with the "reciprocity" feature of the gift-exchange contract can be more effective than either of the two contracts alone. In a labor market in which firms choose which of the three contracts to offer, the new hybrid contract yields the highest employee effort, the highest employee payoff, and the highest social welfare (total firm and employee payoff). In addition, although firms choose to offer the three contracts with similar frequency in the first period, they quickly shift to offering the hybrid contract far more often than either of the other contracts in subsequent periods, such that by the last period, 95 percent of firms offer the hybrid contract.

Overall, our results suggest that optimal contracts designed using standard agency analysis are not as effective as the underlying theory suggests, and that incorporating preferences for reciprocity into the design yields more effective contracts. These results are important for accounting researchers and corporate managers because most employment contracts make payments to employees contingent on performance as measured by the accounting system. Understanding how performance measures (e.g., output or firm profit) in combination with other factors (e.g., wage levels) interact with employees' preference for reciprocity to affect their behavior is crucial for successful contract design. We believe this is a rich area of study that can help management accounting researchers better understand and describe aspects of contracts observed in current practice and can help corporate managers design better employment contracts and control systems.

II. EXPERIMENT 1

Theory and Hypothesis

To illustrate the important features of an optimal agency contract versus a gift-exchange contract, we adopt the experimental setting used in several previous gift-exchange studies (Fehr et al. 1993; Fehr et al. 1998; Hannan et al. 2002; Hannan 2005). We use this setting in our experiment for two main reasons. First, it allows us to compare our results directly to the earlier results because we replicate parts of the earlier studies before extending our investigation to address our new research questions. Second, the setting is simple enough to yield an agency theory prediction, which we use as a benchmark for examining the profit-maximizing effectiveness of the optimal agency contract.[1]

Our experimental setting is a single-period agency setting, in which a firm hires an employee to provide effort in production. Employee effort is not observable, and therefore not contractible. Firm profit is observable to both the firm and the employee. The firm's profit (G) and employee's net utility (U) are:

$$G = (q - w)e$$

$$U = w - c(e)$$

---

[1] An agency theory prediction is possible because there is no state uncertainty and no individual reputation effects. As will be shown later, the lack of state uncertainty means our results cannot be explained by participants' risk preferences, and the single-period setting means that participants cannot develop individual reputation in the labor market. Later in the paper we also discuss the possibility that "group" versus "individual" reputation effects come into play in our experiments as participants gain experience.

where:

q = an exogenously given constant representing revenue to the firm;
w = the firm's wage offer;
e = the employee's effort; and
$c(e)$ = the cost of effort, which increases with the level of effort.

The firm's profit function, the employee's net utility function, and the employee's set of possible effort levels and the associated costs are common knowledge. Under standard agency theory, the optimal incentive contract for this setting solves the following problem for the firm:

$$\text{Maximize } [(q - w)e]$$

$$w, e$$

subject to:

$$w - c(e) \geq U_0$$

$$w - c(e) \geq w - c(e')$$

where:

U_0 = the employee's reservation net utility;
$e, e' \in E$ and $e \neq e'$ (E = the employee's set of possible effort levels).

The solution to this problem is well known. Since the link between employee effort and firm profit is not affected by any external uncertainty, the firm is able to perfectly infer the employee's effort by observing firm profit. Therefore, the firm should use the following "forcing" contract (Harris and Raviv 1978):

$$w = w_H = c(e^*) + U_0 \text{ if } G = (q - w_H)e^*; \text{ and } w = w_L, \text{ otherwise}$$

where e^* = the effort level that maximizes firm profit and $w_L < w_H$.

Under this contract, a rational employee will always choose e^* because it is the only incentive-compatible effort choice. Thus, the firm achieves the first-best outcome by simply paying the market-clearing wage (w_H).

Previous gift-exchange studies that used the setting described above were silent regarding whether firm profit was contractible. If firm profit is contractible, then the contract used in the previous gift-exchange experiments would be theoretically suboptimal. That gift-exchange contract allows firms to offer any wage they wish within the specified range, and then the employees who accept a wage offer can choose any effort level they wish within the specified range. Under this contract, the conventional subgame perfect equilibrium prediction is that, regardless of the firm's wage offer, the employee will always choose the lowest effort possible to maximize his net payoff. Anticipating this, the firm always offers the lowest wage possible.

In contrast to the conventional agency theory prediction, the experiments that tested the gift-exchange contract described above (Fehr et al. 1993; Fehr et al. 1998; Hannan et al. 2002; Hannan 2005) found that both firms' wage offers and employees' effort levels were higher than the subgame perfect equilibrium predictions, and that effort levels were positively associated with wage offers. In addition, firm profit generated by employees' reciprocal effort was, on average, higher than if the firm had paid only the subgame perfect equilibrium wage. These results provided support for Akerlof's (1982) gift-exchange theory.[2] However, such evidence does not fully address the economic implications of reciprocity because the gift-exchange contract used in these studies is a theoretically suboptimal contract. That is, if firm profit is contractible, the theoretically optimal agency contract is the forcing contract described earlier, and by definition, this optimal contract should yield higher firm profit than the suboptimal gift-exchange contract. The hypothesis we test in Experiment 1 is based on this standard economic reasoning:

> **H1:** The theoretically optimal contract for a specific setting will produce higher firm profit than a gift-exchange contract in that same setting.

Experimental Setting and Contracts

The market setting used in Experiment 1 and in our subsequent experiments is the one described earlier. Adopting the parameters used by Hannan et al. (2002) and Hannan (2005), the firm's profit function is:[3]

$$\textit{Firm profit} = (120 - \textit{wage}) \times \textit{employee effort} \tag{1}$$

where $\textit{wage} \in \{20, 21, \ldots, 120\}$ and $\textit{employee effort} \in \{0.1, 0.2, \ldots, 1.0\}$.

The employee's net payoff is:

$$\textit{Employee's net payoff} = \textit{wage} - \textit{cost of effort} \tag{2}$$

where the cost of effort to the employee is as follows:

Effort	0.1	0.2	0.3	0.4	0.5	0.6	0.7	0.8	0.9	1.0
Cost	0	1	2	4	6	8	10	12	15	18

Firm profit is observable to both the firm and the employee. The employee's effort level is not directly observable by the firm, and therefore not contractible. However, because the link between employee effort and firm profit is not affected by any external uncertainty, the firm can perfectly infer the employee's effort by observing firm profit. Also, because there is no external uncertainty, risk preferences do not affect firms' or employees' behavior.

---

[2] Gift-exchange theory has also been supported by field and archival research (Raff and Summers 1987; Blinder and Choi 1990; Agell and Lundborg 1995; Campbell and Kamlani 1997; Huang et al. 1998; Bewley 1999; Bellemare and Shearer 2007). However, to control for potential confounds in the field, experiments have proven to be a very useful tool to test the predictive validity of the gift-exchange model.

[3] In all of our experiments, all amounts were expressed in "lira," an experimental currency. At the end of each experiment, participants' earnings were converted to dollars at the rate of 50 lira = $1.00. The experiments took between approximately 35 minutes (Experiment 1) and an hour (Experiments 2 and 3) to complete, and participants' average earnings were as follows: Experiment 1 (firms/employees, $13.08/$14.05); Experiment 2 (firms/employees, $13.03/$13.83); Experiment 3 (firms/employees, $18.29/$16.28).

For this setting, the optimal contract (hereafter, OPT contract) is determined by solving the maximization problem identified earlier using the parameters shown above. This yields the following optimal contract:[4]

$$\begin{aligned} \textit{wage}\ (w_H) &= 38 \text{ if firm profit is 82 (i.e., employee chooses effort level 1);} \\ \textit{wage}\ (w_L) &= 15 \text{ if firm profit is not 82 (i.e., employee chooses any other} \\ &\quad \text{effort level).} \end{aligned} \tag{3}$$

Note that the firm's profit function in Equation (1) above can be rewritten as:

$$\textit{Firm profit} = (\textit{120} \times \textit{employee effort}) - (\textit{wage} \times \textit{employee effort}). \tag{4}$$

In Equation (4), the first product (120 × employee effort) represents the output produced by the employee. Therefore, using Equation (4), the optimal agency contract in Equation (3) above can be rewritten as:

$$\begin{aligned} \textit{wage}\ (w_H) &= 38 \text{ if output is 120 (i.e., employee chooses effort level 1);} \\ \textit{wage}\ (w_L) &= 15 \text{ if output is not 120 (i.e., employee chooses any other} \\ &\quad \text{effort level).} \end{aligned} \tag{5}$$

We used this output-based version of the optimal contract, i.e., Equation (5) above, in all of our experiments because, as will become clear in Experiment 3, this allowed us to hold constant the description of the optimal contract in the instructions for all of our experiments. Faced with the optimal contract, rational employees will always choose effort level 1 and earn a payoff of 20 (wage of 38 − the cost of effort of 18), generating firm profit of 82 (output of 120 − wage of 38 × effort of 1).[5]

---

[4] The maximization problem is solved using Grossman and Hart's (1983) approach, in which the firm's problem is to choose the contract that maximizes firm profit from among the contracts that implement every possible effort level with the minimum cost to the firm. The contract resulting from applying this approach offers a high wage (w_H) of 38 for firm profit of 82 and a low wage (w_L) for any firm profit other than 82. The low wage (w_L) was set at 15 to ensure that employees were always financially better off by choosing the theoretically optimal effort level of 1 and to avoid negative payoffs for employees who chose a suboptimal effort level. The first reason required that w_L be below 20 because if it were above 20 (e.g., 21), the employee could choose a low effort level (e.g., 0.1) and earn a higher net payoff (wage of 21 − cost of effort of 0 = 21) than the net payoff that they would earn by choosing the optimal effort level of 1 (wage of 38 − cost of effort of 18 = 20). If w_L were 20, then employees would be indifferent between choosing the optimal effort of 1 and the lowest effort of 0.1. The second reason requires that w_L be at least 15. If w_L were below 15 (e.g., 14), then employees who chose the suboptimal effort level of 0.9 would have earned a negative payoff of −1 (wage of 14 − cost of effort of 15 = −1). We avoided negative payoffs because research suggests that loss aversion (Tversky and Kahneman 1991) can bias experimental results when participants face losses and because negative payoffs are prohibited by the institutional body that approves research with human subjects. Given these constraints, we could have set w_L at any wage between 15 and 19. We chose 15 because this allowed firms to impose the highest penalty on employees who did not choose the optimal effort level of 1.

In contrast to the OPT contract, the gift-exchange contract (hereafter, GE contract) for the basic setting described above allows the firm to offer any wage between 20 and 120 and the employee to respond with any effort level between 0.1 and 1. This contract is theoretically suboptimal because agency theory predicts that rational employees will always choose the lowest effort level (0.1) and, anticipating this, the firm will always offer the lowest wage (20). Thus, firm profit will be 10 under the GE contract [(revenue of 120 − wage of 20) × effort of 0.1], which is lower than the firm profit of 82 achieved under the OPT contract. The employee's theoretical payoff under the GE contract is 20, the same as that under the OPT contract.[6]

Design and Procedures

There are two between-subject contract conditions in Experiment 1: GE contract and OPT contract. In the GE contract condition, the GE contract is the only contract available in the market. In the OPT contract condition, the OPT contract is the only contract available in the market.

Participants in Experiment 1 were 80 M.B.A. students, with an average of 5.5 years of full-time work experience. Four experimental sessions were conducted for each of the two contract conditions. In each experimental session, ten participants were randomly assigned the role of firm or employee (five in each role), and they retained this role throughout the session.[7] The data from the four sessions for each contract condition were pooled since there were no significant differences across sessions. Therefore, the final data set contains 20 firms and 20 employees in each of the two contract conditions.

In each session, firms and employees interacted for six periods. To ensure anonymity, firms and employees were identified only by a firm number and an employee number, respectively. To preclude individual reputation formation, new firm (employee) numbers were assigned to each firm (employee) at the start of each of the six periods. Firm and employee participants were seated in the same room, but were divided by a screen. The screen prevented firm and employee participants from seeing each other, but allowed all participants to see the experimenter and a blackboard in the front of the room.

At the start of each session, the experimental instructions were read aloud to all participants. After an overview of the experimental setting and task was provided, the specifics of the contract for that session (either OPT or GE) were explained in detail (see details provided earlier in the "Experimental Setting and Contracts" section). Throughout the session, participants in either of the contract conditions were unaware of the existence of the other contract. Terms such as "gift-exchange" and "optimal" were not used in the experiment to describe the contracts to avoid biasing participants' behavior. Instead, the contract used in any session was simply referred to as "the employment contract." A spreadsheet showing both firm profit and the employee's net payoff for every possible wage-effort combination under the contract used for that session was provided to all participants for use throughout the session. Before the start of the first period of each session, participants completed a series of exercises to ensure that they fully understood the setting, their contract, and how to read the payoff spreadsheet.

---

[5] There is a minor parametric difference in the employee's payoff function between this study and the previous studies by Hannan et al. (2002) and Hannan (2005). The previous studies impose a fixed cost of 20 on employees, making their reservation net utility zero. In this study, the fixed cost is eliminated, making the reservation net utility 20. This change avoids the situation where, under the optimal agency contract, employees receive a zero net payoff even if they choose the firm's desired effort level of 1 (wage of 38 − cost of effort of 18 − fixed cost of 20 = 0), providing employees with no financial motivation to choose the desired effort level. In the current study, employees receive a positive payoff (wage of 38 − cost of effort of 18 = 20) if they choose the desired effort level of 1, providing them with a monetary motivation to do so. This parametric change has no effect on the predictions of interest regarding firms' or employees' behavior in this study as compared to the previous studies.

[6] Although the nominal amount of the market-clearing wage differs between the OPT contract (38) and the GE contract (20), both contracts pay employees the same net payoff of 20. That is, the employee's payoff net of the cost of effort is identical under the OPT contract (wage of 38 − cost of effort of 18 = 20) and the GE contract (wage of 20 − cost of effort of 0 = 20).

[7] Labor markets with an equal number of firms and employees were used in our experiments for two reasons: First, prior research shows that the market's supply-and-demand condition has little impact on the gift-exchange relation (Fehr et al. 1998; Brandts and Charness 2004). Second, although some labor markets may have excess supply or excess demand, examining a balanced market provides a benchmark for possible future studies.

For the GE contract sessions, each of the six periods consisted of the steps in the timeline shown in Panel A of Figure 1. In Steps 1 and 2, firms decided whether to enter the labor market for that period, and those that did, decided on a wage offer and recorded it on their personal record sheet (used to determine their payoff for the period) and a communication form (used to communicate information to employees). Firms that chose not to enter the labor market received zero profit for that period.[8] Next, communication forms were collected from firm participants and all wage offers and associated firm numbers were posted on the blackboard in the front of the room. In Step 3, the employee numbers randomly assigned for that period determined the order in which employees could accept a wage offer (i.e., employee #1 had the first opportunity; employee #2 had the second opportunity, etc.). When an employee's turn came, he could accept any one of the posted wage offers that was still available by silently raising the number of fingers representing the associated firm number, in which case the communication form from the firm offering that wage was given to him. Alternatively, he could choose not to accept any wage offer, in which case his payoff for that period was zero. Once a specific firm's wage offer was accepted, it was no longer available to any other employee.

In Step 4, each employee who accepted a wage offer in Step 3 chose an effort level and recorded it on the communication form, which was then passed back to the firm participant so that he could calculate his profit (Step 5).[9] Employees calculated their payoffs (Step 5) after recording their effort choice on their personal record sheet. The period ended when the experimenter collected the completed personal record sheets from both firms and employees. The procedures in Steps 1–5 described above were repeated for each of the six periods.

The procedures for the OPT contract sessions were the same as for the GE contract shown in Panel A of Figure 1, except that in Step 2, the OPT contract was the only contract that the firm could offer. Because wage offers were fixed and dichotomous under the OPT contract (i.e., the wage is 38 if the output is 120, otherwise the wage is 15), firms did not need to decide on a wage offer in Step 2. The experimenter simply posted on the blackboard the identification numbers of the firms that chose to enter the labor market (all offering the same OPT contract and wages), and, in Step 3, employees chose either to accept a specific firm's contract or to remain unemployed and earn zero for that period.

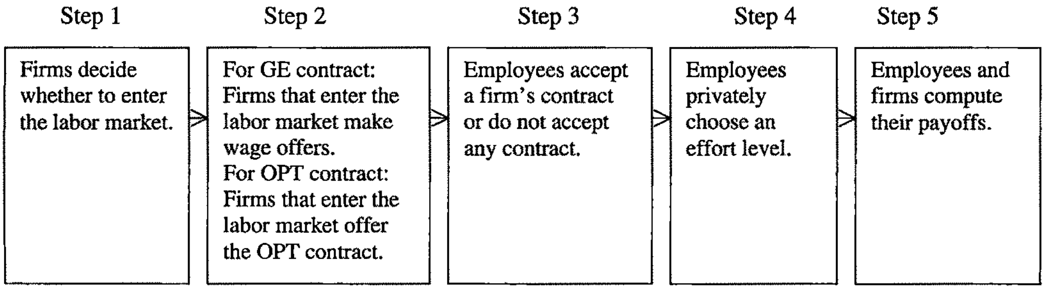
---

[8] Firms almost always chose to enter the labor market (97 percent of the time). Multiple-comparison tests using logistic regression show that the choice to enter the market does not differ significantly across experiments (ps ≥ 0.20, two-tailed). The rare choice of a firm to not enter the market appears to be a reaction to an employee's choice of effort level 0.1 (the lowest effort) or contract rejection in the preceding period.

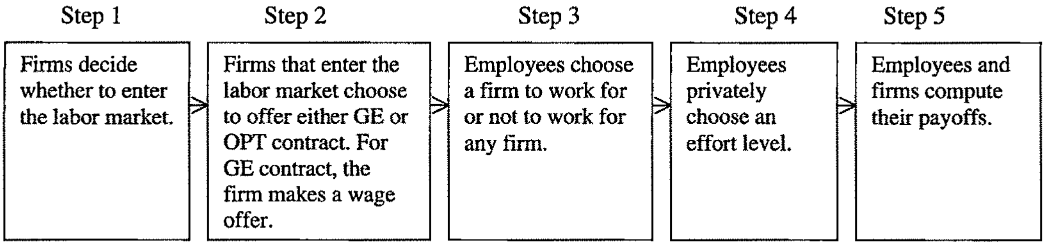
[9] We allowed firm participants to learn the employee's effort choice simply for purposes of calculating the profit they earned in the experiment. We retained the common assumption of non-contractibility of effort from standard agency theory by not allowing firm participants to condition their wage offers on effort in our experiments.

FIGURE 1
Experimental Timelines

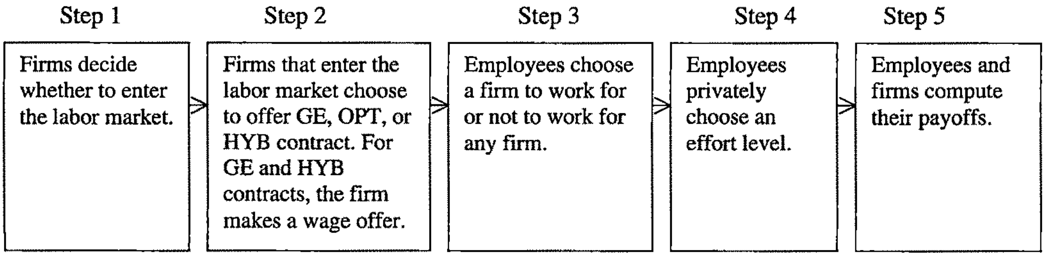
Panel A: Timeline of Experiment 1


Step 1
Firms decide whether to enter the labor market.
Step 2
For GE contract: Firms that enter the labor market make wage offers.
For OPT contract: Firms that enter the labor market offer the OPT contract.
Step 3
Employees accept a firm's contract or do not accept any contract.
Step 4
Employees privately choose an effort level.
Step 5
Employees and firms compute their payoffs.

Panel B: Timeline of Experiment 2


Step 1
Firms decide whether to enter the labor market.
Step 2
Firms that enter the labor market choose to offer either GE or OPT contract. For GE contract, the firm makes a wage offer.
Step 3
Employees choose a firm to work for or not to work for any firm.
Step 4
Employees privately choose an effort level.
Step 5
Employees and firms compute their payoffs.

Panel C: Timeline of Experiment 3


Step 1
Firms decide whether to enter the labor market.
Step 2
Firms that enter the labor market choose to offer GE, OPT, or HYB contract. For GE and HYB contracts, the firm makes a wage offer.
Step 3
Employees choose a firm to work for or not to work for any firm.
Step 4
Employees privately choose an effort level.
Step 5
Employees and firms compute their payoffs.

Steps 4 and 5 for the OPT contract were identical to those described above for the GE contract sessions. Steps 1–5 for the OPT contract were repeated for each of the six periods.

Results of Experiment 1

Panel A of Table 1 provides descriptive statistics for the results of Experiment 1. For each contract, the table reports average firm profit, firms' average wage offer, employees' average effort level, the percentage of effort level 1 responses (highest effort level), the percentage of effort level 0.1 responses (lowest effort level), the percentage of contract rejections, average employee payoff, and average total wealth (firm profit + employee payoff). The highest effort level (1) represents the optimal effort level under the OPT

TABLE 1
Descriptive Statistics for Experiments 1, 2, and 3

Panel A: Descriptive Statistics for Experiment 1 (separate markets for the GE and OPT contracts)

	Periods 1–6		Agency Theory Prediction	
Contract	**GE (n = 114)**	**OPT (n = 119)**	**GE**	**OPT**
Average Firm Profit	23	63	10	82
Average Wage Offer	58	38	20	38
Average Effort	0.41	0.74	0.1	1
Percentage of Effort 1	5%	66%	0%	100%
Percentage of Effort 0.1	26%	21%	100%	0%
Percentage of Rejection	3%	0%	0%	0%
Average Employee Payoff	53	17	20	20
Average Total Wealth	76	81	30	102

Panel B: Descriptive Statistics for Experiment 2 (GE and OPT contracts available in the market)

	Periods 1–6		Periods 7–12		All 12 Periods	
Contract	**GE (n = 57)**	**OPT (n = 59)**	**GE (n = 67)**	**OPT (n = 47)**	**GE (n = 124)**	**OPT (n = 106)**
Average Firm Profit	15	33	17	24	16	29
Average Wage Offer	46	38	35	38	40	38
Average Effort	0.21	0.39	0.22	0.28	0.21	0.34
Percentage of Effort 1	4%	36%	3%	26%	3%	31%
Percentage of Effort 0.1	70%	32%	57%	23%	63%	28%
Percentage of Rejection	0%	32%	6%	49%	3%	40%
Average Employee Payoff	43	12	32	8	37	10
Average Total Wealth	57	44	49	31	53	38

(*continued on next page*)

TABLE 1 (continued)

Panel C: Descriptive Statistics for Experiment 3 (GE, OPT, and HYB contracts available in the market)

	Periods 1–6			Periods 7–12			All 12 Periods		
Contract	**GE (n = 18)**	**OPT (n = 34)**	**HYB (n = 68)**	**GE (n = 4)**	**OPT (n = 10)**	**HYB (n = 102)**	**GE (n = 22)**	**OPT (n = 44)**	**HYB (n = 170)**
Average Firm Profit	14	43	44	11	46	50	14	43	48
Average Wage Offer	61	38	58	45	38	58	58	38	58
Average Effort	0.27	0.51	0.77	0.15	0.55	0.83	0.25	0.52	0.81
Percentage of Effort 1	0%	47%	75%	0%	50%	81%	0%	48%	79%
Percentage of Effort 0.1	67%	38%	19%	50%	50%	14%	64%	41%	16%
Percentage of Rejection	0%	15%	4%	25%	0%	4%	5%	11%	4%
Average Employee Payoff	58	15	38	31	18	37	52	16	38
Average Total Wealth	72	58	82	40	64	87	65	59	85

In Panel B, the agency theory predictions for all reported variables for GE and OPT contracts are identical to those in Panel A. In Panel C, the agency theory predictions for all reported variables for GE and OPT contracts are identical to those in Panel A, and the agency theory predictions for all reported variables for the HYB contract are identical to those for the OPT contract.

Variable Definitions:

Average Firm Profit = mean profit for firms who offered the indicated contract;
Average Wage Offer = mean wage that firms offered in the indicated contract. For OPT and HYB contracts, we only report the high wage that the employee would receive if effort level 1 was chosen. If effort was lower than 1, the wage would be fixed at 15;
Average Effort = mean effort level elicited by the indicated contract;
Percentage of Effort 1 = (the number of the indicated contract that elicited effort level 1)/(the total number of the indicated contract);
Percentage of Effort 0.1 = (the number of the indicated contract that elicited effort level 0.1)/(the total number of the indicated contract);
Percentage of Rejection = (the number of the indicated contract that was rejected)/(the total number of the indicated contract);
Average Employee Payoff = mean net payoff of employees who were offered the indicated contract; and
Average Total Wealth = mean of combined firm profit and employee net payoff under each contract offer for the indicated contract.

contract. The lowest effort level (0.1) represents the theoretically predicted effort level for the GE contract. The lowest effort level also represents the amount of effort an employee would be most likely to choose if he wanted to punish a firm for offering the OPT contract, because this choice imposes the largest penalty on the firm at the lowest cost to the employee. Panel A of Table 1 also reports the agency theory predictions for each of the variables described above for both the GE and OPT contracts.

Hypothesis 1 predicts that the OPT contract will yield higher firm profit than the GE contract. Consistent with H1, a regression of firm profit on contract type (i.e., GE or OPT) shows that firm profit is significantly higher ($t = 10.51$; $p < 0.001$, one-tailed) for the OPT contract (63) than for the GE contract (23).[10] Moreover, Panel A of Figure 2 shows that firm profit for the OPT contract exceeds that of the GE contract for all six periods, with little change in firm profit across the six periods for either contract.

Both firms' and employees' behavior under the GE contract is consistent with the results from prior gift-exchange studies in that, as shown in Panel A of Table 1, both firms' average wage offer (58) and employees' average effort level (0.41) are considerably higher than the conventional subgame perfect equilibrium predictions for the GE contract (wage = 20 and effort = 0.1). In addition, a Tobit regression of effort on wage yields a significant positive association ($z = 3.26$; $p = 0.001$, two-tailed), suggesting that firms and employees exchanged gifts.[11] The effect of the observed gift exchange on firm profit is depicted in Panel A of Figure 2, where we see that firm profit is higher than the theoretical prediction of 10 for the GE contract in all periods. Although these findings replicate prior evidence of reciprocity between firms and employees under the GE contract, the OPT contract still produces higher average firm profit (63) than the GE contract (23) because, under the OPT contract, a substantial percentage of employee effort choices (66 percent) are effort level 1, as predicted by standard agency analysis. However, it is important to note that while these high-effort choices yield higher profit under the OPT, the remaining 34 percent of employee effort choices under the OPT are inconsistent with the standard agency prediction. In particular, as shown in Panel A of Table 1, 21 percent of employee effort choices are the lowest possible level (0.1). This result is consistent with employees reacting negatively to the OPT and responding by imposing the maximum punishment on the firm at the lowest cost to themselves.[12] The combined effect of all suboptimal effort choices on firm profit under the OPT contract is depicted in Panel A of Figure 2, where we see that firm profit does not reach the theoretical prediction of 82 in any period.

In summary, the results of Experiment 1 demonstrate that, consistent with the standard agency theory prediction, the OPT contract yields higher firm profit than the GE contract. However, despite this predicted and observed difference in firm profit, the OPT contract yields less firm profit and the GE contract yields more firm profit than standard agency theory predicts. Hence, employees react negatively to the nonreciprocal aspects of the OPT contract and positively to the reciprocal aspects of the GE contract, but these effects do not dominate the economic forces that underlie the standard agency theory prediction that the OPT contract will yield more firm profit than the GE contract.

---

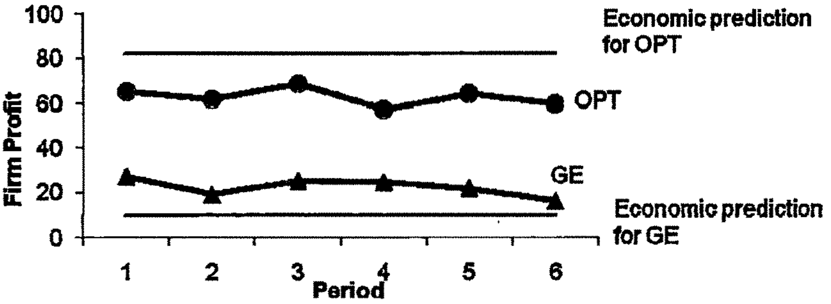
[10] In this study, all regression analyses that involve within-subject observations include the subject as a cluster variable to control for repeated measures. That is, our regression models treat each subject's responses as a cluster, and estimates robust standard errors after adjusting for intra-cluster correlations (Williams 2000; Wooldridge 2002).

[11] The Tobit model controls for the fact that the dependent measure, effort, is a censored value.

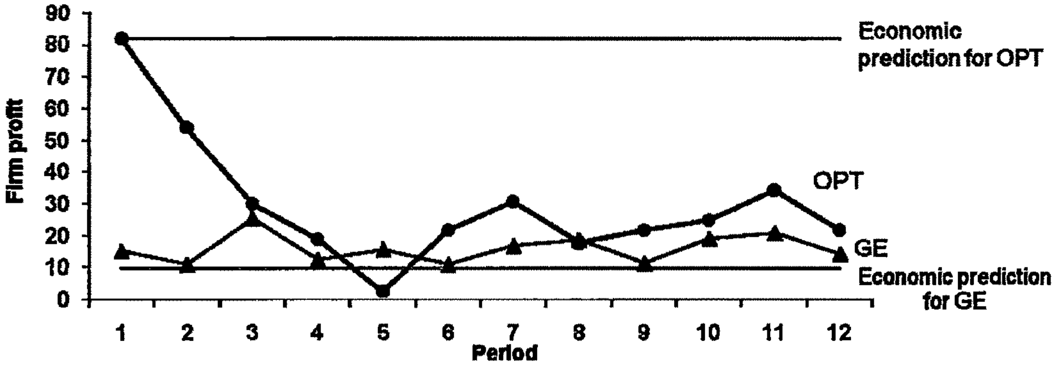
[12] There is no obvious explanation for the small portion (13 percent) of effort choices that fall between 0.1 and 1.

FIGURE 2
Average Firm Profit by Contract and Period

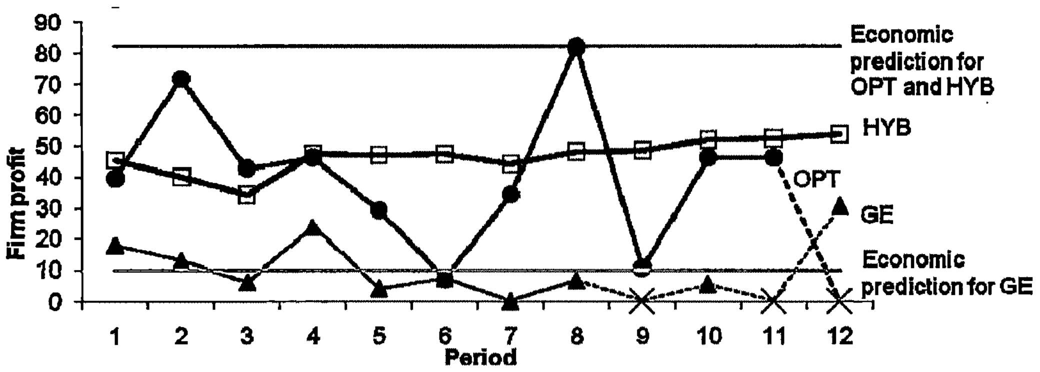
Panel A: Experiment 1



Panel B: Experiment 2



Panel C: Experiment 3



In Panel C, the dotted lines and the "x"s on the horizontal axis for Periods 9, 11, and 12 indicate that the contract was not offered in that period.

III. EXPERIMENT 2

Theory and Hypotheses

As indicated earlier, the optimal agency contract is designed to pay the lowest wage necessary to "force" employees via incentives to select the effort level that maximizes firm profit. Although this is a critical feature of the optimal agency contract, the contract's nonreciprocal nature may be less salient to employees when, as in Experiment 1, no alternative contract is available for comparison. On the other hand, in a market in which both a reciprocity-based gift-exchange contract and the optimal agency contract are available, employees are more likely to view the optimal contract as nonreciprocal. Research suggests that employees generally expect the firm to maintain a reciprocal relationship, and react negatively (e.g., shirk, resign) if the firm is perceived to have violated this relationship (Rousseau 1989; Robinson et al. 1994; Robinson and Rousseau 1994). In our setting, if a firm chooses to offer the OPT contract when the GE contract is available, we hypothesize that employees are more likely to punish the firm by rejecting the OPT contract entirely or by accepting it and providing lower effort, even when this is costly to employees. Such actions by employees would reduce firm profit. This leads to the first hypothesis we test in Experiment 2:

> **H2:** When firms offer the OPT contract in markets in which the GE contract is also available, firm profit will be lower than when firms offer the OPT contract in markets in which only the OPT contract is available.

In markets in which both the OPT and GE contracts are available, employees are likely to gain a better understanding of the payoff implications of each contract as they experience these payoffs. This reasoning suggests that experience with the contracts will enable employees to see more clearly the nonreciprocal nature of the OPT contract versus the reciprocal nature of the GE contract, making them more likely to punish firms that offer the OPT contract. Thus, firm profit would decrease under the OPT contract as employees gain experience, which, in turn, would cause the difference in firm profit between the two contracts to decrease with experience. In the extreme case, the difference in firm profit could even reverse, such that the GE contract would produce higher firm profit than the OPT contract. Because there is no clear basis on which to predict which of the two contracts will yield higher firm profit as employees gain experience with the contracts, we do not offer a directional hypothesis, but rather pose the issue as a two-tailed research question.

> **Research Question:** Will the OPT or GE contract yield higher firm profit as employees gain experience with both contracts in a market in which firms choose which contract to offer?

When deciding which contract to offer, firms will likely consider employees' reaction to earlier contract offers. While our study uses anonymity and random re-matching of participants each period to prevent individual reputation formation, it is still possible that firm participants either hold prior beliefs about how employees will react to their offers or will form beliefs about how employees will react as they observe how employees as a group react to different offers in earlier periods (i.e., a type of "group" reputation formation as the experimental session progresses).[13] That is, some firms may anticipate that employees will react negatively to the OPT contract, and therefore begin by offering the GE contract in the first period. Other firms may start off by offering the OPT contract in hopes of maximizing profits, switching to the GE contract in later periods if this strategy proves ineffective. Such behavior would be consistent with Shields and Waller's (1988) finding that firms are adaptive in revising contracts based on previous outcomes. Similarly, Kagel et al. (1996) find that, in a repeated one-shot ultimatum game, proposers who make low offers are punished by responders (i.e., their offers are rejected), and therefore are "forced" to make higher offers. Likewise, in our setting, firms that initially offer the OPT contract may learn from experience that employees react negatively and therefore switch to offering the GE contract. This leads to the second hypothesis we test in Experiment 2:

> **H3:** When given a choice between offering the OPT contract or the GE contract, firms will offer the GE contract more often as they gain experience.

Design and Procedures

Experiment 2 uses a labor market in which firms could choose to offer either the GE contract or the OPT contract from Experiment 1. Participants were 40 M.B.A. students, with an average of 5.2 years of full-time work experience. Four experimental sessions were conducted. As in Experiment 1, in each session, ten participants were randomly assigned the role of firm or employee (five in each role), retaining this role throughout the session. The data from the four sessions were pooled since there were no significant differences across sessions. Therefore, the final data set contains 20 firms and 20 employees. Because we expected employees and firms to change their behavior as they gained experience with the contracts, participants interacted for 12 periods in each session rather than for six periods as in Experiment 1.

Panel B of Figure 1 provides a timeline of the steps in Experiment 2. The procedures are the same as for Experiment 1, except that in Step 2, the firm chooses which of the two contracts (GE or OPT) to offer. After a firm offers a contract, all other procedures for the GE and OPT contracts are the same as for Experiment 1.

Results of Experiment 2

Descriptive data for Experiment 2 are reported in Panel B of Table 1 in the same format as for Experiment 1. Because H2 predicts that, when firms offer the OPT contract when both contracts are available, firm profit will be lower than when the OPT contract is the only contract available, we test this hypothesis by using data from both Experiment 1 and Experiment 2. We compare firm profit for the OPT contract in Experiment 1 (when it is the only contract available; Panel A of Table 1) with that for the OPT contract in periods

---

[13] In studies with anonymous interaction such as ours, it is standard to view the single-period equilibrium prediction as the economic benchmark against which observed behavior is compared. Conventional economic reputation models typically address "individual" reputation formation and require repeated play between players who know the identity of their opponents. However, recently Healy (2007) offered a model that attempts to capture the notion of "group" reputation (see also Kandori 1992). Healy's (2007) model relies on the assumption that a nontrivial percentage of employees are reciprocal and that it is common knowledge that firms believe employee types are correlated (i.e., there is stereotyping). Because reciprocity and stereotyping fall outside standard economic assumptions, Healy (2007, 1752) refers to his model as "a 'mixed' model in which reciprocity and repeated-game effects operate to generate cooperative outcomes."

1–6 of Experiment 2 (when both contracts are available; Panel B of Table 1).[14] Consistent with H2, a regression of firm profit on market condition (Experiment 1 versus Experiment 2) shows that firm profit is significantly lower ($t = 4.86$; $p < 0.001$, one-tailed) for the OPT contract when both contracts are available (33 from Panel A of Table 1) than when only the OPT contract is available (63 from Panel B of Table 1). Firm profit is lower when both contracts are available because employees choose the theoretically optimal effort level (1) significantly less often (Fisher's exact test: $p < 0.001$, one-tailed) when both contracts are available (36 percent) than when only the OPT contract is available (66 percent), and reject the OPT contract significantly more often (Fisher's exact test: $p < 0.001$, one-tailed) when both contracts are available (32 percent) than when only the OPT contract is available (0 percent). These results are consistent with the reasoning underlying H2 that employees react more negatively to the OPT contract if firms choose to offer it when they could have offered the reciprocity-based GE contract.

Our research question for Experiment 2 asks whether the GE or OPT contract produces higher firm profit as employees and firms gain experience with these contracts. We begin by regressing firm profit on contract type (GE or OPT) for Periods 1–6 of Experiment 2. The result indicates that firm profit is higher ($t = 2.83$; $p = 0.01$, two-tailed) under the OPT contract (33) than under the GE contract (15). However, for periods 7–12, regressing firm profit on contract type indicates that firm profit no longer significantly differs ($t = 0.69$; $p = 0.50$, two-tailed) between the OPT contract (24) and the GE Contract (17). As shown in Panel B of Figure 2, firm profit is similar across the two contracts beginning with Period 3. Consistent with this observation, by-period regressions of firm profit on contract type show that the OPT contract produces significantly higher firm profit than the GE contract in Periods 1 and 2 only. Thus, we find that, as participants gain experience with the contracts, the firm profit produced by the OPT contract is statistically indistinguishable from firm profit produced by the reciprocity-based GE contract.

Hypothesis 3 predicts that firms will offer the GE contract more often as they gain experience. Panel A of Figure 3 reports the percentage of firms that offer the GE and OPT contracts for Periods 1–6, 7–12, and 1–12 in Experiment 2, and provides a graph of the percentages of firms that offer GE and OPT contracts by period. The increase in the percentage of GE contracts from Periods 1–6 (49 percent) to Periods 7–12 (59 percent) is marginally significant (Fisher's exact test: $p = 0.09$, one-tailed). In addition, a logistic regression of firms' contract choices on period shows that, over all 12 periods, firms are marginally more likely ($p = 0.07$, one-tailed) to offer the GE contract in later periods. These results provide modest support for H3.

Further analyses examine whether the shift toward offering the GE contract in later periods is consistent with employees reacting negatively to the OPT contract. Consistent with this explanation, we find that the more often employees respond to the OPT contract by providing a suboptimal level of effort (i.e., effort levels lower than 1), the less likely firms are to offer the OPT contract ($t = 2.27$; $p = 0.04$, two-tailed).[15] This result provides

---

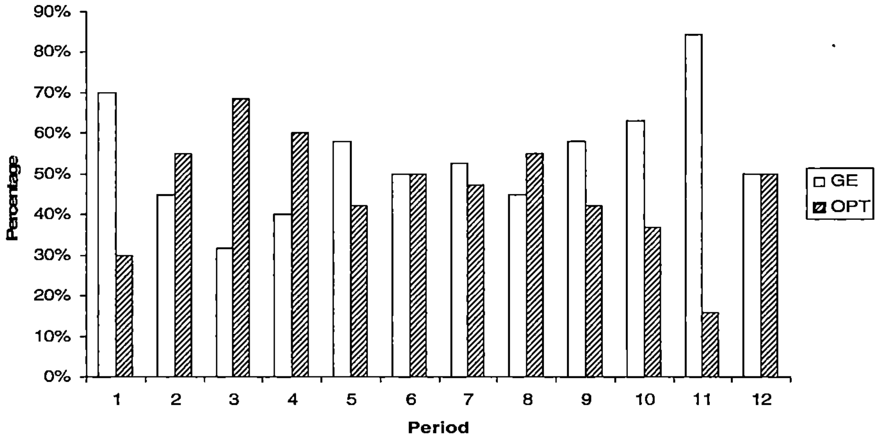
[14] Experiment 2 was conducted immediately after Experiment 1 using a separate set of participants recruited from the same population as those used in Experiment 1. This provides assurance that any differences in results across the two experiments are due to the experimental treatments rather than participant-pool or timing differences.

[15] The reported result is from a regression of the percentage of a firm's OPT contract offers over the 12 periods on: (1) the percentage of OPT contracts offered by the firm that elicited an effort level lower than 1, and (2) the average profit earned by the firm when it offered the GE contract. The latter variable is included as a control variable because the more a firm earns under the GE contract, the more willing it might be to offer the GE contract. Controlling for this possibility allows us to isolate the effect of employees' negative reaction on the firms' contract offers.

FIGURE 3
Frequency of Contract Offers

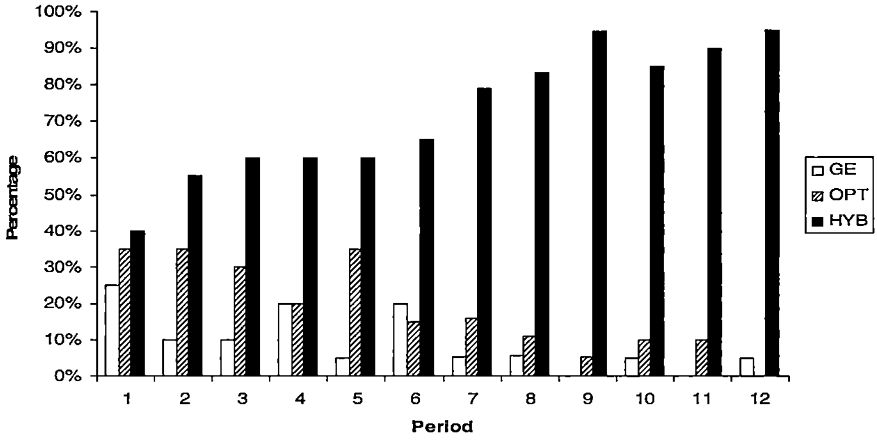
Panel A: Percentage of Contract Offers by Period for Experiment 2

Contract	Periods 1–6	Periods 7–12	Periods 1–12
GE	49%	59%	54%
OPT	51%	41%	46%



Panel B: Percentage of Contract Offers by Period for Experiment 3

Contract	Periods 1–6	Periods 7–12	Periods 1–12
GE	15%	3%	9%
OPT	28%	9%	19%
HYB	57%	88%	72%



evidence consistent with our interpretation that firms begin to shift to offering the GE contract in response to being punished for offering the OPT contract.

One additional interesting aspect of Experiment 2 is that the reason that firm profit becomes statistically indistinguishable between the OPT and GE contract within only a few periods is that employees react more negatively to the OPT contract with experience rather than because employees react more positively to the GE contract with experience. This pattern is clearly depicted in Panel B of Figure 2, which shows that firm profit for the OPT begins at the economic prediction of 82, but drops rapidly to nearly the same level as that for the GE contract, whereas firm profit for the GE contract begins slightly above the economic prediction of 10 and does not change much with experience.

IV. EXPERIMENT 3

Theory and Research Question

The results of Experiment 2 show that employees react more negatively to the OPT contract when firms choose to offer it when they could have offered the more reciprocal GE contract. Nevertheless, the OPT contract still yields firm profit that is statistically indistinguishable from firm profit for the GE contract. The reason for this result is that the OPT contract is able to "force" some, but not all, employees to provide high effort, which, in turn, produces higher firm profit. Even so, the OPT contract performs less effectively than standard agency theory predicts. On the other hand, the GE contract yields slightly more firm profit than standard agency analysis predicts because employees provide unenforceable reciprocal effort in exchange for theoretically unpredicted high wage offers from firms. This raises the question of whether a contract that combines the most powerful features of both contracts might be more effective than either contract alone.

A powerful feature of the OPT contract is its "forcing" feature. Employees are forced to provide the effort level that maximizes the firm's profit because this is the only way they can maximize their own payoffs. A powerful feature of the GE contract is its "reciprocity" feature. Employees apparently prefer to work under a contract that allows firms to elicit reciprocal effort by offering a higher wage. More importantly, employees reciprocate with higher effort when firms offer a higher wage. Thus, both the "forcing" feature of the optimal contract and the "reciprocity" feature of the gift-exchange contract can help firms increase profit.

The question we address in Experiment 3 is whether an alternative hybrid contract that combines both the "forcing" feature of the OPT contract and the "reciprocity" feature of the GE contract can be more effective than either the OPT contract or the GE contract alone. If this were the case, then we would expect that firms would offer this new hybrid contract more often than either of the other contracts, especially as firms and employees gain experience.

Design and Procedures

The design of Experiment 3 is the same as for Experiment 2, except that in Experiment 3, firms are allowed to offer one of three contracts: (1) the OPT contract, (2) the GE contract, or (3) a new hybrid contract (hereafter, HYB contract) that combines the "forcing" feature of the OPT contract with the "reciprocity" feature of the GE contract. Incorporating the forcing feature in the HYB contract means that, as with the OPT contract, employees can only receive the high wage (now set by the firm) if they produce output of 120 (i.e., the output generated by providing the optimal effort level of 1). Otherwise, they receive the low wage (15).[16] Incorporating the reciprocity feature in the HYB contract means that firms can now set the high wage that employees earn if the employees produce output of 120 (generated by the optimal effort level of 1). That is, under the HYB contract, firms can choose to set the high wage above the fixed optimal level of 38 for the OPT contract (possible wage offers range from 20 to 120), and thereby try to induce reciprocal effort from employees. Despite the ability of the firm participants to set the wage above 38, the agency theory predictions for the HYB contract are identical to those for the OPT contract. That is, the firm should always set the wage at the optimal wage of 38 because any amount above that is suboptimal, as it is unnecessary to induce the optimal level of effort.

Participants were 40 M.B.A. students, with an average of 4.2 years of full-time work experience. Four experimental sessions were conducted. As in Experiments 1 and 2, in each session, ten participants were randomly assigned the role of firm or employee (five in each role), retaining this role throughout the session. The data from the four sessions were pooled since there were no significant differences across sessions. Therefore, the final data set contains 20 firms and 20 employees.

A timeline of the steps for Experiment 3 is shown in Panel C of Figure 1. All steps are the same as in Experiment 2 (see Panel B of Figure 1), except that in Step 2, firms can choose one of three contracts as explained above. Procedures for the GE and OPT contract are the same as in Experiments 1 and 2. Procedures for the HYB contract are the same as for the OPT contract, except that, as is the case for the GE contract, firms make a wage offer in Step 2.

Results of Experiment 3

Panel C of Table 1 reports descriptive statistics for Experiment 3. Panel C of Figure 2 and Panel B of Figure 3 depict the average firm profit and frequency of contract offers, respectively. Our main research question for Experiment 3 is whether the new HYB contract can generate as much or more firm profit than the OPT or GE contracts. We begin our analysis by comparing average firm profit for the three contracts across all 12 periods. Regressions of firm profit on contract type show that firm profit for the HYB contract (48) is statistically indistinguishable ($t = 0.45$; $p = 0.66$, two-tailed) from that of the OPT contract (43); whereas firm profit for the GE contract (14) is significantly below (ps ≤ 0.001, two-tailed) that of either of the other two contracts. Because firm profit for the GE contract is well below that of the OPT and HYB contracts, the rest of our analysis focuses only on the latter two contracts.

Despite the similarity in average firm profit between the HYB and OPT contracts, there are significant differences on other important measures. As shown in Panel C of Figure 2, firm profit is considerably more variable for the OPT contract than for the HYB contract (Levene's robust variance test: $p < 0.001$, two-tailed). Also, wage offers are significantly higher ($t = 9.57$; $p < 0.001$, two-tailed) for the HYB contract (58) than for the OPT contract (fixed at 38), and this difference in wage offers is associated with significantly higher employee effort ($t = 3.82$; $p < 0.001$, two-tailed) for the HYB contract (0.81) than for the OPT contract (0.52).[17] Consistent with the higher effort under the HYB contract than under the OPT contract, the percentage of lowest effort level (0.1) choices (16 percent) and contract rejection (4 percent) for the HYB contract (20 percent combined) are significantly less (Fisher's exact test: $p < 0.001$, two-tailed) than for the OPT contract (41 percent + 11 percent = 52 percent combined). As explained more fully below, these differences are consistent with employees punishing firms for offering the OPT contract and rewarding firms for offering the more reciprocal HYB contract.[18]

---

[16] In Experiment 3, under both the HYB and OPT contracts, employees were informed that the high wage would be paid only if the output was 120, just as they were for the OPT contract in the first two experiments (see Equation (5) and related discussion in Section II). Such consistency in wording across the OPT and HYB contracts would have been impossible if the high wage had been conditioned on firm profit rather than output. That is, under the HYB contract the firm profit depends on both the output and the wage the firm chooses to offer, so the contract could not be based on a fixed firm profit of 82 as it could be under the OPT contract. As explained in Section II, conditioning the high wage on output is economically equivalent to conditioning it on firm profit.

The most striking difference between the HYB and OPT contracts is the percentage of times that firms choose to offer each of them. As shown in Panel B of Figure 3, across all 12 periods, firms choose to offer the HYB contract 72 percent of the time, while offering the OPT contract only 19 percent of the time. Moreover, in the first period, the HYB and OPT contracts are offered with similar frequency (40 percent and 35 percent of contract offers, respectively), but the spread between the frequency of these two contract offers increases rapidly, such that by the last period, 95 percent of the firms offer the HYB contract.

To understand why virtually all firms offer the HYB rather than the OPT contract as they gain experience, we need to return to the results for employee effort and firm profit. In order for the OPT contract to produce the theoretically predicted firm profit of 82, employees must provide the theoretically optimal effort level (1) every time they face the OPT contract. In contrast to this prediction, employees facing the OPT contract only provide the optimal effort level 48 percent of the time. In the other 52 percent of the cases, employees either provide the lowest effort level 0.1 or reject the contract, resulting in average firm profit of only 43. Compare these results to those for the HYB contract, in which employees choose the highest effort level (1) 79 percent of the time, which, when combined with the smaller number of times that employees provide a lower effort level or reject the contract (21 percent combined), results in average firm profit of 48.

The results reported above are consistent with firms experiencing employees' negative reaction to the OPT contract and relatively positive reaction to the HYB contract, and responding by shifting to offering the HYB contract. The negative reaction is evidenced by the lower effort and related decrease in firm profit for the OPT contract. The positive reaction is evidenced by the fact that, under the HYB contract, employees reciprocate by offering the highest effort level (1) more often when firms offer a wage higher than the optimal wage. The average wage offered by firms under the HYB contract (58) is significantly higher ($t = 16.13$, $p < 0.001$) than the optimal wage (38), and wage offers are positively related to effort levels ($z = 4.42$, $p < 0.001$, two-tailed).[19] It is important to note that firms do not make themselves worse off by shifting to the HYB contract, because, as reported earlier, firm profit under the HYB contract is, if anything, higher than, albeit statistically indistinguishable from, firm profit under the OPT contract.

---

[17] The reported results are, respectively, from a regression of wage on contract type (OPT or HYB) and a regression of effort on contract type (OPT or HYB).

[18] An analysis of the order in which employees accept the HYB and OPT contract offers provides further evidence that employees prefer the HYB contract to the OPT contract. There are a total of 48 stand-alone markets in Experiment 3 (four sessions with 12 periods in each session). At least one firm (often most firms) choose to offer the HYB contract in all 48 markets, and the HYB contract is the first contract accepted by an employee in 39 of these 48 markets. Compare this to the OPT contract, which is offered by at least one firm in only 23 of the 48 markets, and which is the first contract accepted by an employee in only one of those 23 markets. In fact, the OPT contract is the last contract accepted in 22 of the 23 cases. These results show that employees strongly prefer the HYB contract to the OPT contract.

[19] The reported association represents the result of an ordered logistic regression of effort level on wage, where effort level is a dummy variable coded as 1 for effort level 1, and 0 otherwise. Effort level is dichotomized because virtually all (99 percent) of effort choices for accepted HYB contracts are either the highest level (1) or lowest level (0.1).

A final observation that helps explain why the HYB contract dominates the market in Experiment 3 is that it yields higher social welfare. As can be seen in Panel C of Table 1, the total wealth generated by the HYB contract (85) is higher ($t = 2.55$; $p = 0.02$, two-tailed) than that generated by the OPT (59).[20] This is primarily due to the significantly higher ($t = 10.9$; $p < 0.001$, two-tailed) employee payoffs for the HYB contract (38) than for the OPT contract (16).[21]

In summary, it appears that what is happening in the labor market in Experiment 3 is that firms learn that employees' negative reaction to their OPT offers cause their profit to drop substantially. In response, they switch to the HYB contract and offer a wage that is higher than the theoretically optimal wage of 38, which results in enough reciprocal employee effort to yield firm profit at the same level as the OPT contract. Thus, the market settles into a pattern of HYB contract offers that produce welfare gains for employees, while simultaneously producing firm profit at least as high as under the OPT contract.

V. DISCUSSION AND CONCLUSION

We report the results of three labor market experiments designed to examine the effectiveness of an optimal agency contract as compared to two alternative contracts. The first alternative contract is a reciprocity-based gift-exchange contract used in a series of previous studies. For example, Hannan (2005) used this contract in her study of the effects of wages and firm profit on employee effort, and demonstrated, among other things, that reciprocity between firms and employees can yield higher firm profit than standard agency analysis predicts. Moreover, the broader accounting literature suggests that it may be possible to take advantage of this reciprocity norm to motivate employee effort (Sprinkle 2003).

However, since prior studies did not compare the results from the gift-exchange contract used in those studies with results from the agency contract that would be optimal for the setting used in those studies, the implications of these previous studies for management accounting research are uncertain. That is, from an agency theory perspective, firms should maximize profit by adopting the optimal contract and, therefore, would not need to rely on gift exchange. Our study addresses this issue by examining both contracts. In our first experiment, we find that, consistent with standard agency analysis, the optimal agency contract yields higher firm profit than the gift-exchange contract in markets in which the only contract available is either the optimal contract or the gift-exchange contract. Conversely, in our second experiment, when firms can choose which of the two contracts to offer, employees punish firms that offer the optimal contract by reducing effort or by rejecting the contract. This results in lower firm profit, especially as firms and employees gain experience with the contracts, such that in later periods of the experiment, firm profit under the optimal contract is statistically indistinguishable from that under the gift-exchange contract.

In our third experiment, we include a hybrid contract that combines the "forcing" feature of the optimal contract with the "reciprocity" feature of the gift-exchange contract. In a market in which firms can choose whether to offer the optimal contract, the gift-exchange contract, or the hybrid contract, we find that the hybrid contract dominates the market as firms and employees gain experience with the alternative contracts. Specifically, firms rarely offer the gift-exchange contract, and employees reduce their effort or reject the contract entirely when they are offered the optimal contract. Consequently, firms quickly switch to the hybrid contract rather than the optimal contract, such that in later periods of the experiment, virtually all firms offer the hybrid contract. The hybrid contract yields the highest employee effort, the highest employee payoff, at least as much firm profit as the optimal contract, and the highest total social welfare (combined firm and employee payoff).

---

[20] The reported result is from a regression of the total wealth on contract type (OPT or HYB).

[21] The reported result is from a regression of employee net payoff on contract type (OPT or HYB).

Overall, our results suggest that optimal contracts designed using standard agency analysis are likely to be less effective than the theory suggests, and that incorporating preferences for reciprocity can improve contract design. In the spirit of attempting to explain existing employment practices, we believe that many incentive systems currently used in practice already broadly reflect both of the effort-inducing (and profit-maximizing) features we incorporated into our hybrid contract. This strikes us as very reasonable, as firms would be unlikely to offer a pure gift-exchange contract because the power of such contracts to generate effort and firm profit relies entirely on trusting employees to reciprocate if the firm pays a wage above the market-clearing level. Likewise, it is also unlikely that firms would rely exclusively on financial incentives to force employees to provide effort as is the case in optimal agency contracts, because they most likely understand that employees may react negatively to such contracts. Hence, rather than rely exclusively on either financial incentives or reciprocity alone to induce effort, firms' compensation policies most likely try to incorporate both features. For example, Hansen (2007) explains that in determining employees' compensation, Microsoft Corporation uses a stringent, forced-distribution rating system to evaluate employee performance (a type of "forcing" feature), but then pays employees rated as top performers significantly more than other employees (a "reciprocity" feature).

Notwithstanding this example, there may not be a large number of contracts that *explicitly* include both the "forcing" and "reciprocity" features of the hybrid contract examined in our third experiment. Nevertheless, we believe that the general characteristics of these features are typically part of the broader *implicit* compensation policies applied in practice (Rousseau 1989, 1995). For instance, while many firms may not explicitly indicate the level of performance that triggers an employee's termination, there is still almost surely some implicit threat of being fired for poor performance. Similarly, while there may not be an explicit statement regarding the amount of effort that a firm expects from employees who have been awarded larger than expected pay raises, there is almost surely some implied or informally understood expectation that, in exchange for a large pay raise, employees will increase their effort, or at least maintain effort at an already high level. Thus, we believe that both the "forcing" and "reciprocity" features included in our hybrid contract reflect important aspects of implicit compensation policies observed in practice.

Several limitations of this study could be addressed in future research. For example, our experiments used a one-shot setting to preclude individual reputation formation. However, in practice firms and employees often have an ongoing relationship over multiple periods. Now that baseline results have been established for the case without individual reputation, future research could examine the incremental effects of individual reputation arising from repeated play with fixed firm-employee pairings. Also, future research could investigate the behavior of firms and employees in settings in which employees' effort choices cannot be inferred with certainty by the firm. Since we now have baseline results for the case without external uncertainty, it could be informative to examine the extent to which reciprocity affects behavior in settings with state uncertainty.

REFERENCES

Agell, J., and P. Lundborg. 1995. Theories of pay and unemployment: Survey evidence from Swedish manufacturing firms. *The Scandinavian Journal of Economics* 97 (2): 295–307.

Akerlof, G. A. 1982. Labor contracts as partial gift exchange. *The Quarterly Journal of Economics* 97 (4): 543–569.

Baiman, S. 1982. Agency research in managerial accounting: A survey. *Journal of Accounting Literature* 1: 154–213.

Bellemare, C., and B. Shearer. 2007. Gift exchange within a firm: Evidence from a field experiment. Working paper, Institute for the Study of Labor.

Bewley, T. F. 1999. *Why Wages Don't Fall During a Recession.* Cambridge, MA: Harvard University Press.

Blinder, A. S., and D. H. Choi. 1990. A shred of evidence on theories of wage stickiness. *The Quarterly Journal of Economics* 55 (4): 1003–1015.

Brandts, J., and G. Charness. 2004. Do labor market conditions affect gift-exchange? Some experimental evidence. *The Economic Journal* 114 (497): 684–708.

Campbell, C. M., and K. S. Kamlani. 1997. The reasons for wage rigidity: Evidence from a survey of firms. *The Quarterly Journal of Economics* 112 (3): 759–789.

Fehr, E., G. Kirchsteiger, and A. Riedl. 1993. Does fairness prevent market clearing? An experimental investigation. *The Quarterly Journal of Economics* 108 (2): 437–459.

———, E. Kirchler, A. Weichbold, and S. Gachter. 1998. When social norms overpower competition: Gift exchange in experimental labor market. *Journal of Labor Economics* 16 (2): 324–351.

Grossman, S. J., and O. D. Hart. 1983. An analysis of the principal-agent problem. *Econometrica* 51 (1): 7–46.

Hannan, R. L., J. H. Kagel, and D. V. Moser. 2002. Partial gift exchange in experimental labor market: Impact of subject population differences, productivity differences, and effort requests on behavior. *Journal of Labor Economics* 20 (4): 923–951.

———. 2005. The combined effect of wages and firm profit on employee effort. *The Accounting Review* 80 (1): 167–188.

Hansen, F. 2007. Compensation and salary forecast: Lackluster performance. *Workforce Management* (November 7): 39–45.

Harris, M., and A. Raviv. 1978. Some results on incentive contracts with applications to education and employment, health insurance, and law enforcement. *The American Economic Review* 68 (1): 20–30.

Healy, P. J. 2007. Group reputations, stereotypes, and cooperation in a repeated labor market. *The American Economic Review* 97 (5): 1751–1773.

Huang, T., A. H. H. Hallam, P. F. Orazem, and E. M. Paterno. 1998. Empirical tests of efficiency wage models. *Economica* 65 (257): 125–143.

Indjejikian, R. J. 1999. Performance evaluation and compensation research: An agency perspective. *Accounting Horizons* 13 (2): 147–157.

Kagel, J. H., C. K. Kim, and D. V. Moser. 1996. Fairness in ultimatum games with asymmetric information and asymmetric payoffs. *Games and Economic Behavior* 13: 100–110.

Kandori, M. 1992. Social norms and community enforcement. *The Review of Economic Studies* 59 (1): 63–80.

Lambert, R. A. 2001. Contracting theory and accounting. *Journal of Accounting and Economics* 32: 3–87.

Raff, D. M. G., and L. H. Summers. 1987. Did Henry Ford pay efficiency wages? *Journal of Labor Economics* 5 (4): S57–S86.

Robinson, S. L., M. S. Kraatz, and D. M. Rousseau. 1994. Changing obligations and the psychological contract: A longitudinal study. *Academy of Management Journal* 37 (1): 137–152.

———, and D. M. Rousseau. 1994. Violating the psychological contract: Not the exception but the norm. *Journal of Organizational Behavior* 15: 245–259.

Rousseau, D. M. 1989. Psychological and implied contracts in organizations. *Employee Responsibilities and Rights Journal* 2 (2): 121–139.

———. 1995. *Psychological Contracts in Organizations: Understanding Written and Unwritten Agreements*. Thousand Oaks, CA: Sage.

Shields, M. D., and W. S. Waller. 1988. A behavioral study of accounting variables in performance-incentive contracts. *Accounting, Organizations and Society* 13 (6): 581–594.

Sprinkle, G. B. 2003. Perspectives on experimental research in managerial accounting. *Accounting, Organizations and Society* 28 (2–3): 287–318.

Tversky, A., and D. Kahneman. 1991. Loss aversion in riskless choice: A reference dependent model. *The Quarterly Journal of Economics* 106 (4): 1039–1062.

Williams, R. L. 2000. A note on robust variance estimation for cluster-correlated data. *Biometrics* 56 (2): 645–646.

Wooldridge, J. M. 2002. *Econometric Analysis of Cross Section and Panel Data*. Cambridge, MA: The MIT Press.

THE ACCOUNTING REVIEW
Vol. 84, No. 5
2009
pp. 1695–1697

American Accounting Association
DOI: 10.2308/accr.2009.84.5.1695

BOOK REVIEW

Stephen A. Zeff, Editor

Editor's note: Two copies of books for review should be sent to the Book Review Editor: Stephen A. Zeff, Rice University, Jesse H. Jones Graduate School of Business, 6100 Main St., Houston, TX 77005. The policy of *The Accounting Review* is to publish only those reviews solicited by the Book Review Editor. Unsolicited reviews will not be accepted.

MARTIN BLOOM, *Double Accounting for Goodwill: A Problem Redefined* (London, U.K.: Routledge, 2008, ISBN 13:978-0-415-43748-6, pp. ix, 231).

The definition, measurement, and accounting for goodwill have been a controversial issue in financial accounting over the past century. Many different approaches have been adopted in official accounting standards; however, none has been universally accepted or lasted for long. Each has been replaced by an equally unsatisfactory treatment. The underlying reason for the mess surrounding goodwill accounting is aptly summarized by Professor Robert Sterling as quoted on page 1: "We accountants do not resolve issues, we abandon them. ... We debate them loud and long ... until another issue comes along. ... The explanation for our inability to resolve issues is to be found in the way we conceive issues. We conceive of the issues in such a way that they are in principle irresolvable ..."

Bloom has responded to this challenge in his book, and he has developed a simple, sound, and useful solution to the problem of accounting for goodwill. He explains and critically examines each of the approaches adopted over time, and why it was jettisoned in favor of the next one adopted. However, as Sterling observed, none of the solutions adopted solved the problem, because, as formulated, it was not capable of being resolved. In no case did the method adopted provide useful and reliable information to investors, notwithstanding that the reported goodwill asset constituted an increasing proportion of corporate total assets over time, and was the largest asset reported by some. Reliable information on goodwill is critical for shareholders' and potential investors' decisions on share investments and appraisal of corporate performance. Bloom's explanations and analyses of the issues involved in accounting for goodwill are comprehensive, thorough, and incisive. His solution is appropriate, simple, and can be readily implemented to provide the information required by shareholders without distorting the reporting of the entity's ongoing operations.

In his treatise, Bloom begins by examining the concept of goodwill and the various interpretations assigned to it in accounting, whether it should be treated as an asset in the accounts of a company, and how it may be measured. A range of views on each of these issues is presented. He then examines the usual distinction made in accounting between internally generated goodwill and purchased goodwill, and whether the former should be accounted for. In the accounting literature, internally generated goodwill is presumed to be developed by processes and non-specific expenditures within the entity itself. Because these matters are generally difficult to identify, isolate, and measure, such goodwill is not accounted for. However, Bloom points out that such goodwill can also be generated from the firm having some monopolistic powers in the market, enabling it to charge higher prices (quoting the economist Kenneth Boulding), and/or from the corporate management having superior ability in managing operations where future market conditions are uncertain and frequently changing (quoting another economist, Frank H. Knight). Not recognizing internally generated goodwill is one of only two issues on which accountants have agreed.

Bloom examines all the methods that have been used for the treatment of purchased goodwill. It is an identifiable transaction expenditure incurred in buying another business, and traditionally it has been contended that it should be reported as such. The big question is: How to do so? This depends on answers to the following questions:

(1) Does the goodwill have a finite or an everlasting life? If the former, how can the life over which it contributes to revenues be measured? Normally, the decision is an arbitrary one, and limits of 40 years or 20 years have been prescribed by the accounting authorities.
(2) Which amortization method should be used? Normally the straight-line method is used, but some companies prefer the reverse sum-of-the years' digits method, which produces only modest charges in the early years. Bloom notes that, if expenditures are incurred to maintain the level of goodwill, then charging these expenditures plus amortization expense results in double charging.
(3) Should the goodwill amortization expense be charged against revenue, thereby reducing reported profit, or should it be charged against owners' equity as if it were a prepayment on future profits? Normally, the expense charge is preferred.
(4) Should the company use purchase or pooling accounting for the acquisition of another entity? If pooling is used, as occurred for many years in the U.S., the goodwill purchased in the acquisition is not recognized.

Because of the inability to reach long-term agreement on the above questions, each survived for only limited periods. The current "fashionable treatment" (p. 101) is to use the "impairment method." This is now the IASB's prescribed method. Impairment loss is defined as "the amount by which the carrying amount of an asset ... exceeds its recoverable amount" (IAS 36, IASB 2004, para. 6). Advocates of the impairment method had challenged the assumption underlying the amortization approach that purchased goodwill has only a limited life. Rather, they argued that it could be maintained, increased, or dissipated by good or poor management, ongoing expenditures on it, or by market factors. Hence, reductions in goodwill should be recognized only if and when they occur. But again, fundamental questions arise as to:

(1) How to recognize and measure the impairment?
(2) Over what period should the impairment be assessed?
(3) How reliable are the estimates of expected future cash flows underpinning the measure of recoverable value, given the uncertainties surrounding future operations?
(4) What discount rate should be used to measure the present values of expected future cash flows?
(5) Should recoverable value be calculated for the aggregate entity, or separately for each income-generating unit, as is required? There may be synergies across the company's operations.

Finally, it appears to be inconsistent to recognize reductions in assessed goodwill but not improvements, even to reinstate previous reductions when market conditions have improved. Given that the whole process of estimating recoverable values is based on arbitrary and subjective estimates, Bloom does not endorse the impairment solution.

Fortunately, the author's solution is a simple one, *viz.*, to report goodwill outside the balance sheet in a supplementary financial statement, the Market Capitalization Statement (MCS). In the MCS, total entity goodwill is reported at its stock-market determined value (calculated by the share price times the number of shares) less the current realizable value of its net financial assets, tangible assets, and identifiable intangible assets. The measure is a robust one that avoids all the problems encountered in the traditional accounting treatments. The goodwill measure is an aggregate one that does not distinguish between purchased and internally generated goodwill, and does not allocate it to individual parts of the business. Stock exchanges provide an ongoing valuation of the listed companies based on all the publicly available information about each company and the markets in which they operate. Their rules prescribe the release of timely and continuous information for the benefit of investors. A company's share price impounds all this information and summarizes the future profit prospects of each, as seen by large numbers of investors.

The measure of goodwill so determined is important information for shareholders. It indicates investors' expectations of changes in the future earning power of companies. When it is expected to increase, the value of goodwill rises, and so on. It thus provides important information to shareholders on the past and expected profit performance of the company, and thus influences their share investment decisions.

Underlying this approach to reporting goodwill is the old double account system used in nineteenth century Britain by railway and other infrastructure companies. The balance sheet was separated into two parts—a capital section, which reported the capital contributed by shareholders, and long-term borrowings, which funded the operating capacity of the company, and the general balance sheet, which reported the "floating" (i.e., current) assets and liabilities of the company that are in a continuing state of change from daily operations. Bloom illustrates his proposed balance sheet and Market Capitalization Statement using Chambers' system of valuing all assets and liabilities at their net realizable values.

Double accounting for goodwill provides a quality examination of the issues and problems encountered in accounting for goodwill in traditional accounting, and why none of them is satisfactory. Bloom has applied Sterling's dictum in doing so. The proposed solution is a sound and simple one that not only avoids all the previous problems encountered, but, most importantly, also provides essential and important information to investors. The only modification I recommend to the approach advocated is to allow the use of current replacement costs of physical assets for companies as an option in those situations where investment in them is substantial. In fact, Tobin's "q ratio," which Bloom approvingly cites, is based on asset replacement costs. This ratio is that between the market value of a company and the replacement cost of its tangible assets together with the realizable values of its net financial assets. Similarly, this approach is adopted in Edwards and Bell (1961) in their concept of objective goodwill.

REFERENCES

Edwards, E. O., and P. W. Bell. 1961. *The Theory and Measurement of Business Income*. Berkeley, CA: University of California Press.

International Accounting Standards Board (IASB). 2004. *Impairment of Assets*. International Accounting Standard 36. London, U.K.: IASB.

ALLAN BARTON
Emeritus Professor of Accounting
Australian National University

EDITORIAL POLICY AND STYLE INFORMATION

EDITORIAL POLICY

According to the policies set by the Publications Committee (which were endorsed by the Executive Committee and were published in the *Accounting Education News*, June 1987), *The Accounting Review* "should be viewed as the premier journal for publishing articles reporting the results of accounting research and explaining and illustrating related research methodology. The scope of acceptable articles should embrace any research methodology and any accounting-related subject, as long as the articles meet the standards established for publication in the journal ... No special sections should be necessary. The primary, but not exclusive, audience should be—as it is now—academicians, graduate students, and others interested in accounting research."

The primary criterion for publication in *The Accounting Review* is the significance of the contribution an article makes to the literature. Topical areas of interest to the journal include accounting information systems, auditing and assurance services, financial accounting, management accounting, taxation, and all other areas of accounting, broadly defined. The journal is also open to all rigorous research methods.

The efficiency and effectiveness of the editorial review process is critically dependent upon the actions of both authors submitting papers and the reviewers. Authors accept the responsibility of preparing research papers at a level suitable for evaluation by independent reviewers. Such preparation, therefore, should include subjecting the manuscript to critique by colleagues and others and revising it accordingly prior to submission. The review process is not to be used as a means of obtaining feedback at early stages of developing the research.

Reviewers and editors are responsible for providing constructive and prompt evaluations of submitted research papers based on the significance of their contribution and on the rigor of analysis and presentation.

MANUSCRIPT PREPARATION AND STYLE

The Accounting Review's manuscript preparation guidelines follow *The Chicago Manual of Style* (15th ed.; University of Chicago Press). Another helpful guide to usage and style is *The Elements of Style*, by William Strunk, Jr., and E. B. White (Macmillan). Spelling follows *Webster's Collegiate Dictionary*.

FORMAT

1. All manuscripts should be formatted in 12-point font on 8 1/2 x 11" paper and should be double-spaced, except for indented quotations.
2. Manuscripts should be as concise as the subject and research method permit, generally not to exceed 7,000 words.
3. Margins should be at least one inch from top, bottom, and sides.
4. To promote anonymous review, authors should not identify themselves directly or indirectly in their papers or in experimental test instruments included with the submission. Single authors should not use the editorial "we."
5. A cover page should show the title of the paper, all authors' names, titles and affiliations, email addresses, and any acknowledgments.

Pagination: All pages, including tables, appendices and references, should be serially numbered. Major sections should be numbered in Roman numerals. Subsections should not be numbered.

Numbers: Spell out numbers from one to ten, except when used in tables and lists, and when used with mathematical, statistical, scientific, or technical units and quantities, such as distances, weights and measures. For example: *three days; 3 kilometers; 30 years.* All other numbers are expressed numerically.

Percentages and Decimal Fractions: In nontechnical copy use the word percent in the text.

Hyphens: Use a hyphen to join unit modifiers or to clarify usage. For example: *a cross-sectional equation; re-form.* See *Webster's* for correct usage.

Keywords: The abstract must be followed by at least three keywords to assist in indexing the paper and identifying qualified reviewers.

Data Availability: A line immediately following the Keyword identifiers should indicate whether the data are available (see "Data Availability" later in this statement).

ABSTRACT/INTRODUCTION

An Abstract of about 100 words (150 maximum) should be presented on a separate page immediately preceeding the text. The Abstract should concisely inform the reader of the manuscript's topic, its methods, and its findings. Keywords and the Data Availability statements should follow the Abstract. The text of the paper should start with a section labeled "I. Introduction," which provides more details about the paper's purpose, motivation, methodology, and findings. Both the Abstract and the Introduction should be relatively nontechnical, yet clear enough for an informed reader to understand the manuscript's contribution. The manuscript's title, but neither the author's name nor other identification designations, should appear on the Abstract page.

TABLES AND FIGURES

The author should note the following general requirements:

1. Each table and figure (graphic) should appear on a separate page and should be placed at the end of the text. Each should bear an Arabic number and a complete title indicating the exact contents of the table or figure. Tables and figures should define each variable. The titles and definitions should be sufficiently detailed to enable the reader to interpret the tables and figures without reference to the text.
2. A reference to each graphic should be made in the text.
3. The author should indicate where each graphic should be inserted in the text.
4. Graphics should be reasonably interpreted without reference to the text.
5. Source lines and notes should be included as necessary.

Equations: Equations should be numbered in parentheses flush with the right-hand margin.

DOCUMENTATION

Citations: Within-text citations are made using an author-year format. Cited works must correspond to the list of works listed in the "References" section. Authors should make an effort to include the relevant page numbers in the within-text citations.

1. In the text, works are cited as follows: author's last name and year, without comma, in parentheses. For example: one author, (Berry 2003); two authors, (Fehr and Schmidt 2003); three or more authors, (Dechow et al. 1998); more than one work cited, (Cole and Yakushiji 1984; Dechow et al. 1995; Levitt 1998); with two works by the same author(s), (Nelson 2003, 2005).
2. Unless confusion would result, do not use "p." or "pp." before page numbers. For example, (Dechow and Dichev 2002, 41–42).
3. For cited works that include more than one work by an author (or same co-authors) that are published in the same year, the suffix a, b, etc., is to follow the date in the within-text citations and in the "References" section. For example: (Baiman and Rajan 2002a, 2002b).
4. When the author's name is mentioned in the text, it need not be repeated in the citation. For example: "Cohen et al. (2005) provide ..."
5. Citations to institutional works shoulduse acronyms or short titles where practicable. For example: (NCFFR, The Treadway Commission 1987).
6. If the paper refers to statutes, legal treatises, or court cases, citations acceptable in law reviews should be used.

Reference List: Every manuscript must include a "References" section that contains only those works cited within the text. Each entry should contain all information necessary for unambiguous identification of the published work. Use the following formats (which follow *The Chicago Manual of Style*):

1. Arrange citations in *alphabetical order* according to the surname of the first author or the name of the institution or body responsible for the published work.
2. Arrange citations in *chronological order* when two or more works are by the same first author (regardless of co-authors). Two or more works by the *same author(s)* published in the same year are distinguished by letters a, b, etc., after the year.
3. Use authors' initials instead of proper names.
4. For two or more authors, separate authors with a comma, including a comma before "and" (Dechow, P. M., R. Sloan, and A. Sweeney).
5. Date of publication follows the author's (authors') name(s).
6. Titles of journals or newspapers should not be abbreviated.

Sample entries are as follows:

Baiman, S., and M. Rajan. 2002a. The role of information and opportunism in the choice of buyer-supplier relationships. *Journal of Accounting Research* 40 (2): 247–278.

———, and ———. 2002b. Incentive issues in inter-firm relationships. *Accounting, Organizations and Society* 27 (3): 213–238.

Berry, R. 2003. Testimony before the Senate Committee on Homeland Security and Governmental Affairs Permanent Subcommittee on Investigations. November 18. Available at: http://hsgac.senate.gov/_files/111803berry.pdf.

Cohen, D., A. Dey, and T. Lys. 2005. The Sarbanes Oxley Act of 2002: Implications for compensation structure and risk-taking incentives of CEOs. Working paper, New York University, University of Chicago, and Northwestern University.

Cole, R., and T. Yakushiji, eds. 1984. *The American and Japanese Auto Industries in Transition.* Ann Arbor, MI: University of Michigan.

Dechow, P. M., R. Sloan, and A. Sweeney. 1995. Detecting earnings management. *The Accounting Review* 70 (2): 193–225.

———, S. P. Kothari, and R. L. Watts. 1998. The relation between earnings and cash flows. *Journal of Accounting and Economics* 25: 133–168.

———, and I. Dichev. 2002. The quality of accruals and earnings: The role of accrual estimation errors, *The Accounting Review* 77 (Supplement): 35–59.

Easton, P. 2003. Discussion of: The predictive value of expenses excluded from pro forma earnings. *Review of Accounting Studies* 8: 175–183.

Engel, E., R. Hayes, and X. Wang. 2007. The Sarbanes-Oxley Act and firms' going-private decisions. *Journal of Accounting and Economics* (forthcoming).

Financial Accounting Standards Board (FASB). 2006. *Accounting for Uncertainty in Income Taxes, and Interpretation of FASB Statement No. 109.* FASB Interpretation No. 48. Financial Accounting Series. Norwalk, CT: FASB.

Fehr, E., and K. Schmidt. 2003. A theory of fairness, competition, and cooperation. In *Advances in Behavioral Economics*, edited by C. Camerer, G. Loewenstein, and M. Rabin, 271–296. New York, NY: Princeton University Press.

Janis, I. L. 1982. *Groupthink: Psychological Studies of Policy Decisions and Fiascoes.* Boston, MA: Houghton Mifflin.

Levitt, A. 1998. The numbers game. Speech delivered at New York University, Center for Law and Business, September 28.

Maggi, G. 1999. The value of commitment with imperfect observability and private information. *RAND Journal of Economics* (Winter) 30: 555–574.

National Commission on Fraudulent Reporting (the Treadway Commission). 1987. *Report of the National Commission on Fraudulent Financial Reporting.* Washington, D.C.: NCFFR.

Nelson, M. W. 2003. Behavioral evidence on the effects of principles- and rules-based standards. *Accounting Horizons* 17 (1): 91–104.

———. 2005. A review of experimental and archival conflicts-of-interest research in auditing. In *Conflicts of Interest: Challenges and Solutions in Business, Law, Medicine, and Public Policy,* edited by D. A. Moore, D. M. Cain, G. Loewenstein, and M. H. Bazerman. Cambridge, U.K.: Cambridge University Press.

Rigdon, E. E., R. E. Schumacker, and W. Wothke. 1998. A comparative review of interaction and nonlinear modeling. In *Interaction and Nonlinear Effects in Structural Equation Modeling*, edited by R. E. Schumacker, and G. A. Marcoulides, 1–16. Mahwah, NJ: Erlbaum Associates.

Schultz, E., and T. Francis. 2002. Companies profit on workers' deaths through "dead peasants" insurance. *Wall Street Journal* (April 19): 1.

Securities and Exchange Commission (SEC). 2002. *Certification of Disclosure in Companies' Quarterly and Annual Reports.* Release Nos. 33-8124, 34-46427. Washington, D.C.: SEC.

U.S. House of Representatives. 2002. The Sarbanes-Oxley Act of 2002. Public Law 107-204 [H. R. 3763]. Washington, D.C.: Government Printing Office.

Footnotes: Footnotes are not used for documentation. Textual footnotes should be used only for extensions and useful excursions of information that if included in the body of the text might disrupt its continuity. Footnotes should be inserted using the "footnote" or "endnote" feature of the word processing software which will automatically number the footnotes throughout the manuscript with superscript Arabic numerals.

SUBMISSION OF MANUSCRIPTS

Authors should note the following guidelines for submitting manuscripts:

1. Manuscripts currently under consideration by another journal or publisher should not be submitted. The author must state upon submission that the work is not submitted or published elsewhere.
2. For manuscripts reporting on field surveys or experiments: If the additional documentation (e.g., questionnaire, case, interview schedule) is sent as a separate file, then all information that might identify the author(s) must be deleted from the instrument.
3. Manuscripts should be submitted via email as Microsoft Word or PDF file to the Senior Editor, Steven Kachelmeier, at email address tar@mccombs.utexas.edu. Please submit separate files for (1) the manuscript's title page with identifying information (not forwarded to reviewers), (2) the manuscript with title page and all other identifying information removed, and (3) any necessary supplemental files, such as experimental instructions and/or response memoranda on invited revisions. The body or subject line of the email should also indicate the submission fee receipt code from AAA (see step 4 below).
4. A submission fee is required unless explicitly waived by the editor for invited revisions of previous submissions. The submission fee should be paid online (VISA or MasterCard only) at: https://aaahq.org/AAAforms/journals/tarsubmit.cfm. If unable to pay online, a check in U.S. funds, made payable to the American Accounting Association, may be sent to Ms. Mary Capps, Editorial Assistant, *The Accounting Review*, The University of Texas at Austin, McCombs School of Business, Department of Accounting CBA 4M.202, 1 University Station B6400, Austin, TX 78712-0211. The submission fee in U.S. funds is $200.00 for members and $400.00 for nonmembers of the AAA. The submission fee is nonrefundable.
5. Revisions must be submitted within 12 months from the decision letter inviting a revision.

COMMENTS

The Accounting Review welcomes submissions of comments on previous *TAR* articles. Comments on articles previously published in *The Accounting Review* will generally be reviewed by two reviewers, usually including an author of the original article (to assist the editor in evaluating whether the submitted comment represents the prior article accurately) and an additional reviewer who is independent of the original article. If a comment is accepted for publication, the original author will generally be invited to reply. All other editorial requirements, as enumerated above, apply to proposed comments.

POLICY ON REPRODUCTION

An objective of *The Accounting Review* is to promote the wide dissemination of the results of systematic scholarly inquiries into the broad field of accounting.

Permission is hereby granted to reproduce any of the contents of the *Review* for use in courses of instruction, as long as the source and American Accounting Association copyright are indicated in any such reproductions.

Written application must be made to the American Accounting Association, 5717 Bessie Drive, Sarasota, FL 34233-2399, for permission to reproduce any of the contents of the *Review* for use in other than courses of instruction—e.g., inclusion in books of readings or in any other publications intended for general distribution. In consideration for the grant of permission by the *Review* in such instances, the applicant must notify the author(s) in writing of the intended use to be made of each reproduction. Normally, the Review will not assess a charge for the waiver of copyright.

Except where otherwise noted in articles, the copyright interest has been transferred to the American Accounting Association. Where the author(s) has (have) not transferred the copyright to the Association, applicants must seek permission to reproduce (for all purposes) directly from the author(s).

POLICY ON DATA AVAILABILITY

The following policy was adopted by the Executive Committee in its 1989 meeting. "An objective of (*The Accounting Review, Accounting Horizons, Issues in Accounting Education*) is to provide the widest possible dissemination of knowledge based on systematic scholarly inquiries into accounting as a field of professional research, and educational activity. As part of this process, authors are encouraged to make their data available for use by others in extending or replicating results reported in their articles. Authors of articles which report data dependent results should footnote the status of data availability and, when pertinent, this should be accompanied by information on how the data may be obtained."

SEPTEMBER 2009 PLACEMENT ADS

The deadline for free position ads to be included in this section is two months prior to the date of publication. *The Accounting Review* is published in January, March, May, July, September, and November. Position ads in the "Placement Ads" are free with the purchase of a job posting in the AAA Career Center. For more information on how to purchase a job posting in the Career Center and receive your free position ad in *The Accounting Review*, go to the AAA website at http://aaahq.org and click on "Career Center," or call our office at 941-921-7747.

501 SOUTHERN ILLINOIS UNIVERSITY CARBONDALE, School of Accountancy is seeking individuals who have an earned Doctorate in Accounting, administrative experience, and a minimum of five years post-Ph.D. teaching experience at the effective date of employment. Candidates must meet all university, college, and school requirements for a tenured, Full Professor, which include: education, teaching, research, and service. The individual should have a nationally recognized, established research record in premier and high-quality publications such as *The Accounting Review*, *Journal of Accounting Research*, *Contemporary Accounting Research*, or *Accounting, Organizations and Society*. Financial and Audit accounting specialties are preferred, but quality applicants in other areas of accounting will be considered. Qualified candidates are invited to submit a letter of application, resume, and contact information for three references by email to: sobery@cba.siu.edu; or by post mail to: Dr. Julie Sobery, Search Committee Chair, School of Accountancy, Mail Code 4631, College of Business Southern Illinois University Carbondale, 1025 Lincoln Drive, Carbondale, IL 62901: Phone: (618) 453-2289: Fax (618) 453-1411. Applications will be accepted until May 29, 2009 or until position is filled. For additional information go to http://www.siu.edu/~affact/jobs.html. SIUC is an Affirmative Action/Equal Opportunity Employer that strives to enhance its ability to develop a diverse faculty and staff and to increase its potential to serve a diverse student population. All applications are welcomed and encouraged and will receive consideration.

502 SAINT JOSEPH'S UNIVERSITY, Department of Accounting, Haub School of Business invites applications for an anticipated tenure-track opening to begin August 2010. Applicants with a specialty in Financial Accounting and/or Managerial Accounting will be considered at the Assistant Professor rank. The position requires that the doctoral degree in Accounting be completed, or near completion by the date of appointment. CPA is desired. Qualifications include a Ph.D. in Accounting or related field or ABD in Accounting. All applications must be submitted online at: https://jobs.sju.edu. Saint Joseph's University is a private, Catholic, Jesuit institution and expects members of its community to be knowledgeable about its mission and to make a positive contribution to that mission. Saint Joseph's University seeks to recruit, develop, and retain a talented and diverse workforce. AA/EOE M/F/D/V.

503 SAINT JOSEPH'S UNIVERSITY, Department of Accounting, Haub School of Business invites applications for an anticipated tenure-track opening, Professor of Accounting and the Edward G. Sutula Chair in Accounting, to begin August 2010. Applicants with a specialty in Financial Accounting and/or Managerial Accounting will be considered at the Professor rank. The position requires that the doctoral degree in Accounting be completed. CPA is desired. Faculty are expected to teach undergraduate and graduate courses, be actively involved in academic research, and provide service to the profession and the university. Review of applications will begin immediately. Visit https://jobs.sju.edu for more details.

504 MICHIGAN TECHNOLOGICAL UNIVERSITY, School of Business and Economics at is seeking outstanding accounting candidates in all areas of specialization for a Tenured or Tenure-Track Position in Accounting as a Assistant, Associate, or Full Professor for August 2010. Candidates must demonstrate research and teaching excellence for a senior position or potential for a junior position. Senior candidates are expected to provide leadership in research and develop sources of external funding. The successful candidate is expected to develop or continue a vigorous, nationally visible research program and contribute effectively to undergraduate and graduate teaching. Candidates need a Ph.D. in accounting from an AACSB-accredited business school. Candidates should send (1) a letter of application describing qualifications for the position, (2) a current *cv*, (3) a list of references, and one-page summary statement describing (4) career research plans, and (5) career teaching plans. The position will remain open until filled. For full consideration submit materials by October 31, 2009. All inquiries and application materials should be sent electronically to the Accounting Search Committee at: sbe-search@mtu.edu. The School of Business and Economics (http://www.sbe.mtu.edu) is AACSB-accredited, offers competitive salaries, attractive benefits, and a reduced teaching load and start-up funding for junior faculty. Michigan Tech is an ADVANCE institution, and has a Dual Career Assistance Program. Minorities and women are especially encouraged to apply. Michigan Technological University is an Affirmative Action/Equal Opportunity Employer.

505 SAINT LOUIS UNIVERSITY, John Cook School of Business, is a Catholic, Jesuit institution dedicated to student learning, research, health care, and service, and is seeking applicants for a one-year visiting, non-tenure appointment in accounting, beginning Spring 2010. Rank is open. A doctorate in accounting with strong research and teaching skills is required. Salary is dependent on level and qualifications. Teaching interest is open and all areas will be considered. All applications must be made online at: http://jobs.slu.edu. Applicants should submit a cover letter, current *curriculum vitae*, and names of three references. All other correspondence regarding this position can be sent to: Dr. Ananth Seetharaman, Chair of Accounting, John Cook School of Business, 3674 Lindell Blvd., Saint Louis University, St. Louis, MO 63108. Saint Louis University is an Affirmative Action/Equal Opportunity Employer (AA/EOE), and encourages nominations of and application from women and minorities.

506 SEATTLE UNIVERSITY, Albers School of Business and Economics seeks a Visiting Professor of Accounting with a Ph.D. or who is ABD for the 2009/2010 academic year. Possible teaching areas include undergraduate or master's level in the area of financial accounting or auditing. Candidates must be committed to excellence in teaching and scholarly research. The Albers School is accredited by the AACSB International and includes a strong undergraduate business program, Master of Professional of Accounting, M.B.A., and Leadership Executive M.B.A. program. Interested applicants should send their *vita* and teaching evaluations to: Dr. Bruce Koch, Chair, Department of Accounting, Seattle University, Seattle, WA 98122; or Email: kochb@seattleu.edu. Seattle University, founded in 1891, continues a 450-year tradition of Jesuit Catholic higher education. The University's Jesuit Catholic ideals underscore its commitment to the centrality of teaching, learning, and scholarship, of values-based education grounded in the Jesuit and Catholic traditions, of service and social justice, of lifelong learning, and of educating the whole person. Located in the heart of dynamic Seattle, the University enrolls approximately 7,500 undergraduate and graduate students in eight colleges and schools. Students enjoy a university ethos characterized by small classes, individualized faculty attention, a strong sense of community, a commitment to diversity, and an outstanding faculty.

507 SEATTLE UNIVERSITY seeks an Assistant/Associate Professor of Accounting for the 2010/2011 academic year. A Ph.D. in Accounting is required. While applications in all areas will be considered, preference will be given to those with experience teaching undergraduate and/or graduate courses in Financial Accounting. Candidates must be committed to excellence in teaching and scholarly research. Our school is accredited by the AACSB International. Albers's programs include an undergraduate program, Master of Professional of Accounting, M.B.A., and Executive M.B.A. program. Applicants should send their *vita*, samples of research, teaching evaluations, and three letters of reference to: Dr. Bruce Koch, Chair, Department of Accounting, Seattle University, Seattle, WA 98122; or Email: kochb@seattleu.edu. Seattle University, founded in 1891, continues a 450-year tradition of Jesuit Catholic higher education. The University's Jesuit Catholic ideals underscore its commitment to the centrality of teaching, learning, and scholarship, of values-based education grounded in the Jesuit and Catholic traditions, of service and social justice, of lifelong learning, and of educating the whole person. Located in the heart of dynamic Seattle, the University enrolls approximately 7,500 undergraduate and graduate students in eight colleges and schools. Students enjoy a university ethos characterized by small classes, individualized faculty attention, a strong sense of community, a commitment to diversity, and an outstanding faculty. Seattle University is an Equal Opportunity/Affirmative Action Employer.

508 UNIVERSITY OF PUERTO RICO IN SAN JUAN has two Accounting faculty positions and is seeking candidates with a Ph.D. or D.B.A. in Accounting. CPA/CMA preferred. Teaching, research, and service expected. Contact Dr. Jose A. Gonzalez at: gtaboada@uprrp.edu.

509 STATE UNIVERSITY OF NEW YORK (SUNY) AT NEW PALTZ, School of Business invites applications for a tenure-track faculty position in Accounting for Fall 2010. Ph.D. in Accounting is required; ABDs near completion are encouraged to apply. Ability to teach Auditing is a plus. The School of Business offers six B.S. degrees and two M.B.A. programs including an M.B.A. in Public Accountancy. The University emphasizes a balance between teaching and research. The search will continue until position is filled. For additional information, contact Dr. Hadi Salavitabar, Dean, School of Business at: Phone: (845) 257-2930; Email: salavith@newpaltz.edu; or refer to http://www.newpaltz.edu/schoolofbusiness. Please send an application letter, resume, three references, and other appropriate materials to: Dr. Hadi Salavitabar, Dean, School of Business, SUNY at New Paltz, 1 Hawk Drive, New Paltz, New York 12561; Email: salavith@newpaltz.edu. SUNY at New Paltz is an AA/EOE/ADA Employer. Women and minorities are encouraged to apply.

510 UNIVERSITY OF MANITOBA, Department of Accounting & Finance, I.H. Asper School of Business in Winnipeg, Canada, invites applications for a full-time tenure-track position at the rank of Assistant Professor within all areas of Accounting. A Ph.D. or D.B.A. in Accounting completed or near completion is required. Salary is commensurate with qualifications, experience, and research record. The I.H. Asper School of Business has approximately 1,400 undergraduate, 150 M.B.A., and 15 Ph.D. students. The I.H. Asper School of Business is AACSB-accredited, and the accounting program is accredited by the Society of Management Accountants of Canada. The School and the accounting group receive excellent support from the three Canadian accounting organizations: The Institute of Chartered Accountants, The Society of Management Accountants, and the Certified General Accountants. The University of Manitoba is the most prestigious university in the province, situated in the capital city, Winnipeg, which has a population of 670,000 people and offers excellent cultural, sporting, and recreational activities. Please send your application letter, current *curriculum vitae*, contact information of three referees, evidence of teaching effectiveness, unofficial transcripts, a description of research, and sample working papers to: Dr. David A. Stangeland, University of Manitoba, Winnipeg, MB, Canada R3T 5V4; Phone: (204) 474-6477); or Email: d_stangeland@umanitoba.ca. Please see our electronic posting on the AAA website. Application materials will be handled under the "Freedom of Information and Protection of Privacy Act" (Manitoba).

511 UNIVERSITY OF SOUTHERN INDIANA, College of Business (AACSB-accredited) invites applications for a tenure-track position in Accounting beginning August 2010. Candidates seeking the rank of Assistant Professor should have a Ph.D. or D.B.A. in Accounting (or a related discipline), but ABDs will be considered at the rank of Instructor until completion of the degree. Preference will be given to candidates with a primary specialization in Accounting Information Systems, although other areas will be seriously considered. This position offers a range of resources for professional development. The College values continuing self-development in scholarship, teaching, and service consistent with AACSB standards, and collaborative interactions with colleagues. The College will move to a state-of-the-art building, which is currently under construction, in Fall 2010. To apply, please send a letter of application describing your interest in the position, *curriculum vita*, teaching philosophy statement, and evidence of teaching effectiveness (if available) and scholarship to: Mary Spahn, College of Business, University of Southern Indiana, 8600 University Blvd., Evansville, IN 47712. Please contact Ms. Spahn at (812) 465-1979 (or mespahn@usi.edu) if you have any questions or need additional information. An original faculty/staff application form, at least three professional references, and copy of transcripts, will be required to complete the application package. Position is open until filled although priority consideration will be given to applications completed by August 31, 2009.

512 ALBION COLLEGE seeks an exceptional teacher who can work well in a selective liberal arts college with a broad-based undergraduate curriculum. Responsibilities include teaching three undergraduate Financial Accounting and Tax Accounting courses per semester and ongoing scholarship. Tenure-track appointment at Assistant or untenured Associate level beginning August 2010. Ph.D. or D.B.A. in accounting preferred; ABD, or M.B.A. with CPA will be considered. Preference given to applications received by June 30, 2009. Submit application letter, *vita*, teaching evaluations, graduate transcript, three letters of recommendation, and research paper. To apply, see: http://www.albion.edu/hr/ACpositions/AcadAffairs/TenureTrackAccounting2010.asp. Albion College has 1,600 students and is located in a small, culturally diverse community in south-central Michigan within an hour's drive of the University of Michigan, Michigan State, and Western Michigan University. Albion College is an Equal Opportunity Employer committed to diversity as a core institutional value.

513 FUDAN UNIVERSITY, one of the top universities in China, invites applications for Assistant, Associate, and Full Professor of Accounting. We offer competitive salary, research, and housing supports. The applicant should have a Ph.D. degree or ABD in the related field from a prestigious university, be able to teach undergraduate and graduate courses in accounting, and publish at major journals. For Associate or Professor level, applicants should be able to demonstrate excellent teaching and research records. For Assistant Professor level, applicants should be able to demonstrate strong potential for research and teaching. Review of applications will start immediately. Applicants please submit your *cv*, job market papers, and three reference letters to: Ms. Yan Zhou, Faculty Office, School of Management, 670 Guo Shun Road, Shanghai 200433; Phone: 86-21-65642410; Fax: 86-21-65654523; Email: hrfdms@fudan.edu.cn; Website: www.fdms.fudan.edu.cn.

514 SOUTHERN ILLINOIS UNIVERSITY EDWARDSVILLE invites applications for a tenure-track position beginning in August 2010. A Ph.D. in Accounting is required, but ABDs nearing completion will be considered. Our primary needs are in the areas of cost and managerial accounting or taxation but qualified candidates with teaching interests in other areas are encouraged to apply. All candidates must be committed to excellence in teaching and scholarly research. The Department of Accounting offers AACSB-accredited programs at both the undergraduate and master's levels. SIUE is located on a suburban campus of 2,700 acres approximately 25 minutes from downtown St. Louis, MO. Send *vita* to: Michael L. Costigan, Department Chair, Department of Accounting, Campus Box 1104, Southern Illinois University Edwardsville, Edwardsville, IL 62026-1104; Email: mcostig@siue.edu. SIUE is a state university; benefits under state sponsored plans will not be available to holders of F1 or J1 visas. SIUE is an Equal Opportunity/Affirmative Action Employer, and applications from women and minorities are especially encouraged.

515 SUSQUEHANNA UNIVERSITY, Sigmund Weis School of Business seeks candidates at the Assistant or Associate Professor rank for a tenure-track position in financial accounting starting Fall 2010. Candidates should possess an earned doctorate in Accounting; ABDs with strong potential considered. Professional certification desirable. The Sigmund Weis School of Business, accredited by AACSB International, focuses exclusively on undergraduate education: 23 faculty members serve 575 majors. Susquehanna, a selective, residential, liberal arts institution with 2,100 students, is located 50 miles north of Harrisburg and about three hours from Philadelphia, Washington, and New York City. For additional information see: http://www.susqu.edu/human_resources/Employment/AsstProfAcct_2-09.html. Please submit letter of interest, *curriculum vita*, unofficial transcripts, and the names, telephone numbers, and email addresses of three professional references to: Dr. Barbara Woods McElroy, c/o Andrea Popp (popp@susqu.edu), or mail to: Sigmund Weis School of Business, Susquehanna University, 514 University Avenue, Selinsgrove, PA 17870. AA/EOE. Applications from women and minorities are encouraged.

516 UNIVERSITY OF LETHBRIDGE is seeking applications for a tenure-track position in accounting at our Lethbridge campus beginning July 2010 at any rank from persons with research interests in financial accounting and teaching interests in intermediate financial accounting and other upper level financial accounting courses. Applicants should possess a relevant doctorate or be in the final stages of completing their dissertation. Professional certification is preferred. A commitment to teaching excellence and publishing research in quality journals is required. The Faculty of Management consists of approximately 65 faculty members and approximately 2,500 undergraduate students located in the three campuses of Lethbridge, Edmonton, and Calgary. The Faculty has a research-based M.Sc. program in Management and strong international exchange programs. Email your application directly to: management.dean@uleth.ca. Include *cv*, teaching evaluations, evidence of research effectiveness and research objectives, and three reference letters.

517 CORNERSTONE RESEARCH, an economic consulting firm specializing in analyzing economic, financial, accounting, and marketing issues, is interested in hiring full-time and summer associates for its offices in Boston, MA; Los Angeles, CA; Menlo Park, CA; New York, NY; San Francisco, CA; and Washington, D.C. Cornerstone Research provides an interesting and rewarding work environment. Associates work directly with esteemed faculty experts in a distinctive "partnership" that combines the strengths of the academic and business worlds. We seek Ph.D. candidates with the ability to apply academic research to real-world issues, present concise explanations of complex analyses, and independently manage projects and junior staff. The ideal candidate will possess a strong empirical background in Financial, Cost, or Auditing, as well as demonstrate excellent interpersonal skills. We also seek experienced candidates for more senior levels. Compensation commensurate with experience. Submit resume, sample research paper, and cover letter that include a statement describing your interest in economic consulting and your location preferences, resume, and letters of reference at: http://associatecareer.cornerstone.com.

518 NORTHEASTERN UNIVERSITY has an opening for a tenure-track position effective September 2010 at the Assistant or Associate Professor level. We seek candidates who have or expect to have a doctorate in accounting or closely related business field before the start date of September 2010 and who are committed to excellence in teaching and research. Teaching experience in accounting, a strong research program focused on top-tier journals, and some business background are preferred. The accounting group provides a supportive environment where collaboration in teaching and research is encouraged. The position offers competitive compensation and teaching responsibilities that enable a notable level of high-impact research. Rank and salary are dependent on qualifications. Summer funding, graduate assistants, and travel support are available. Accounting courses are offered at both the undergraduate and at the graduate levels (M.B.A., M.S./M.B.A., M.S.A; and M.S.Tax). Please send a cover letter, a *cv*, and two samples of scholarly work to: Dr. Marjorie Platt, m.platt@neu.edu. Northeastern University is an Affirmative Action/Equal Opportunity/Title IX Employer and welcomes applications from minorities, women, and persons with disabilities.

519 UNIVERSITY OF MASSACHUSETTS AMHERST, Isenberg School of Management, the Department of Accounting invites applications for two tenure-track positions as Assistant/Associate Professor of Accounting. Start date is September 2010. All candidates should have a doctoral degree by time of appointment and a strong commitment to scholarly research and excellence in teaching. The department has a continuing commitment to enhance the diversity of its faculty, student body, and curriculum. Salary is competitive. Rank and salary are commensurate with experience and qualifications. Please send *c.v.*, contact information for three references, and copies of research papers or publications to: Professor Pieter Elgers, Chair, Search Committee, Department of Accounting and Information Systems, Isenberg School of Management, University of Massachusetts Amherst, 121 Presidents Drive, Amherst, MA 01003. Review of applications will begin on October 1, 2009 but applications will be reviewed until successful candidates are found. The University of Massachusetts Amherst is an Affirmative Action/Equal Opportunity Employer. Applications from women and minority candidates are especially encouraged.

520 TEXAS A&M INTERNATIONAL UNIVERSITY (TAMIU), Division of International Banking and Finance Studies (DIBFS) invites applicants for multiple tenure/tenure-track positions in accounting at Assistant, Associate, and/or Full Professor ranks. Job duties include teaching, producing scholarly research, and participating in service activities. Preference will be given to candidates who specialize in financial, managerial, tax accounting, and international, but other areas will also be considered. Applicants must have earned an appropriate doctorate from a school accredited by the AACSB at the time of appointment. Doctoral degrees from schools not accredited by AACSB may be considered on a case-by-case basis. ABDs may be appointed for visiting positions at the Instructor or Assistant Professor level. Candidates for the Associate or Full Professor level should show that they warrant appointment at that rank. Positions will remain open until filled. Applicants should send an application letter with supporting documentation, including a current resume, transcripts of all graduate work, and at least three references, electronically, to: Ken Hung, Chair at ken.hung@tamiu.edu, or via postal mail to: DIBFS, TAMIU, 5201 University Boulevard, Laredo, TX 78041. TAMIU is an Equal Opportunity Employer. Women and minorities are encouraged to apply.

521 ILLINOIS STATE UNIVERSITY, Department of Accounting invites applications for a full-time, tenure-track Assistant or Associate Professor position in accounting (contingent on budgetary approval and appropriate funding). The successful applicant will begin August 16, 2010. The Department of Accounting is AACSB-accredited and provides a balanced emphasis on quality teaching, research, and service. Candidates are expected to demonstrate the potential for excellence in teaching and scholarly research. Qualifications include a Ph.D. or D.B.A. in Accounting from an AACSB-accredited school. ABDs with substantial progress and close to graduation will be considered. Professional certification and work experience in accounting are desirable. Teaching responsibilities will be primarily in undergraduate and graduate courses in taxation. The normal teaching load is nine credit hours per semester with a reduced load during the first two years. To ensure full consideration, please send a formal letter of application and current *vita* by September 15, 2009 to: Gerry McKean, Interim Chairperson, Illinois State University, Department of Accounting, Campus Box 5520, Normal, IL 61790-5520; Email: gerry.mckean@ilstu.edu. Illinois State is an Equal Opportunity/Affirmative Action university encouraging diversity.

522 CALIFORNIA STATE UNIVERSITY, LONG BEACH invites applicants for a tenure-track accounting faculty position, Assistant/Associate Professor for Fall 2010. Applicants must hold a Ph.D. or a D.B.A. degree with a major in Accounting and have demonstrated potential for excellence in teaching, research, curriculum development, and service contributions. Professional certification is desirable. Compensation will be competitive and be commensurate with the candidate's qualifications and experience. Applicant screening will begin immediately and will continue until the position is filled. Please send a letter of application and *curriculum vitae* to: Dr. Steven A. Fisher, Chair, Department of Accountancy, California State University, Long Beach, 1250 Bellflower Boulevard, Long Beach, CA 90840-8504; Email: sfisher@csulb.edu. CSULB is an urban university in the California State University system. It is located near the ocean in the City of Long Beach. CSULB is an Affirmative Action/Equal Opportunity Employer and has a strong commitment to the principle of diversity.

523 UNIVERSITY OF BALTIMORE, Merrick School of Business, an AACSB-accredited university at both undergraduate and graduate levels, invites applications for a full-time Assistant Professor in Accounting to begin Fall 2010. All research and teaching interests will be considered but a background in auditing is preferred. Candidates are required to hold a doctoral degree (Ph.D. or D.B.A.) in accounting from an AACSB International-accredited university. ABDs will be considered if completion is expected within the first year of appointment. We are looking for balanced individuals who can provide high-quality teaching and publish in peer-reviewed journals. Application materials must include a detailed cover letter, *curriculum vitae*, recent teaching evaluations, and names and addresses of three professional references that may be contacted. Applications may be directed to: Dr. Lourdes White at: lwhite@ubalt.edu (Ernst & Young Professor of Accounting).

524 ROBERT MORRIS UNIVERSITY is a private university with an enrollment of approximately 5,000 undergraduate and graduate students. The undergraduate and graduate degree programs in our School of Business at are accredited by AACSB. The Department of Accounting and Finance invites applications for an Assistant Professor/Associate Professor of Accounting position. Candidates with teaching and research interests in all areas of accounting will be considered. Candidates should either possess an earned doctorate in accounting or be nearing completion of their degree program. Preference will be given to candidates with a sound record of teaching and research. Candidates should submit a letter of application, *curriculum vitae*, and names and contact information for three references to: business@rmu.edu. Review of applications will begin immediately and continue until the position is filled. For additional information, please contact Dr. Frank Flanegin, Phone: (412) 397-2560; or Email: flanegin@rmu.edu. EOE.

525 ROBERT MORRIS UNIVERSITY, Department of Accounting and Finance invites applications for an Assistant Professor/Associate Professor of Taxation position. Candidates should either possess an earned doctorate in accounting or taxation or be nearing completion of their degree program. Individuals who hold a J.D./L.L.M. or a combination of a J.D. and a graduate degree in accounting and who also possess a strong record of publication will be considered. Preference will be given to candidates with a sound record of teaching and research. RMU's degree programs in the School of Business are accredited by AACSB. Candidates should submit a letter of application, *curriculum vitae*, and names and contact information for three references to: business@rmu.edu. Review of applications will begin immediately and will continue until the position is filled. For additional information, contact Frank Flanegin, Phone: (412) 397-2560; or Email: flanegin@rmu.edu. Visit http://www.rmu.edu for information regarding Robert Morris University. EOE.

526 UNIVERSITY OF TENNESSEE, Knoxville, Department of Accounting and Information Management is accepting applications for a full-time, tenure-track position to begin August 1, 2010. This position is an extension (reopening) of the position advertised last year. Candidates should hold a doctoral degree with a specialization in financial accounting or expect to complete all degree requirements by his or her starting date. An additional teaching or research interest in international accounting or taxation would be desirable. Candidates for Assistant Professor should demonstrate excellent teaching potential and have an academic background that strongly suggests an ability to establish an ongoing program of substantive research. Candidates for Associate Professor should demonstrate excellent achievement in teaching and have a substantive research record that reflects a significant ongoing research program. Please send applications to: Kenneth E. Anderson, Search Committee Chair, Department of Accounting and Information Management, CBA, University of Tennessee, Knoxville, TN 37996-0560; Email: kea@utk.edu. The University of Tennessee is an EOE/AA/Title VI/Title IX/Section 504/ADA/ADEA institution in the provision of its education and employment programs and services. All qualified applicants will receive equal consideration for employment without regard to race, color, national origin, religion, sex, pregnancy, marital status, sexual orientation, gender identity, age, physical or mental disability, or covered veteran status.

527 UNIVERSITY OF SCRANTON, Kania School of Management (AACSB) invites applications for its tenure-track position in Taxation beginning August 2010, at the rank of Assistant Professor. Duties include teaching at the undergraduate and graduate levels, scholarly research activities meeting AACSB-related criteria, and service. The primary teaching responsibilities will be in taxation. The faculty member will also be responsible for administering the University's VITA program. The qualifications for the position include an appropriate earned doctorate or ABD progressing toward completion. The University of Scranton is, by tradition and choice, a Catholic and Jesuit university. Applications (including *vita*, transcripts, and teaching evaluations) may be submitted to Dr. Robyn Lawrence at accounting@scranton.edu. Applications will be considered until the position is filled. The University is committed to developing a diverse faculty, staff, and student body and to modeling an inclusive campus community. The University of Scranton is an EOE/Affirmative Action Employer/Educator. Veterans, minority persons, women, and persons with disabilities are encouraged to apply.

528 UNIVERSITY OF NORTH ALABAMA invites applications for a tenure-track faculty position in Accounting/Auditing beginning August 2010. Candidates should have, or expect to have, an earned doctorate in accounting when employment begins. Candidates must have interest in teaching Auditing, and other area(s) from Financial, Systems, Managerial, Cost, or Governmental. Professional certification, auditing experience, and teaching experience are highly desirable. Candidates must be qualified to teach undergraduate and graduate courses. We are interested in faculty who desire a close working relationship with the professional community. Rank will depend upon qualifications. UNA is located in Florence, AL, a part of the Shoals, on the scenic Tennessee River. The Shoals is the cultural, medical, commercial, and recreational center for the region. Submit an application at: https://jobs.una.edu and attach a cover letter, *vitae*, and reference list (contact information for three references). UNA is an Equal Opportunity employer committed to achieving excellence and strength through diversity.

529 OLD DOMINION UNIVERSITY seeks a well-trained scholar, who will publish in major refereed journals and has a strong commitment to teaching, to join the Accounting Faculty at the Assistant or Associate Professor level starting in the Fall 2010 semester. This position involves performing research and teaching in any area of accounting but preferably in Taxation. Preference will be given to candidates with published articles or who have the ability to publish in leading accounting academic journals and who hold professional certifications, especially the CPA or J.D. Applicants must have a Ph.D. in Accounting or related Business Discipline. Applicants for the Associate Professor position must have a proven track record of published research in internationally recognized refereed accounting academic journals. Send *curriculum vita*, and names, addresses, and telephone numbers of three references to: Dr. Douglas E. Ziegenfuss, Chair, Department of Accounting, Constant Hall 2157, Old Dominion University, Norfolk, VA 23529-0229; or Email: dziegenf@odu.edu. Screening of applicants will begin August 31, 2009 and continue until the position is filled. Old Dominion University is an Affirmative Action/Equal Opportunity institution and requires compliance with the Immigration Reform and Control Act of 1986.

530 UNIVERSITY OF KANSAS, School of Business seeks applicants and nominations for the position of Director, Accounting and Information Systems Area, and Deloitte & Touche Fellow. Candidates should have a record of achievement in teaching and research to warrant a tenured appointment at the rank of Full Professor and evidence of an ability to operate effectively in a participative, collegial environment. Previous administrative experience as Department Chair is highly desirable. Professional certification, relevant business experience, recent involvement in mentoring doctoral students, and familiarity with Information Systems curricula are also desirable. The Director will provide academic leadership in faculty development, program growth, relationships with the professional accounting community at the local, regional, and national levels, and will have budgetary and administrative responsibility for the Accounting and Information Systems Area. The Accounting and Information Systems programs offer degrees at the Bachelors, Master's, and Doctoral levels. The Accounting program is separately accredited by the AACSB. For additional information about the School and its faculty and programs, go to: http://www.business.ku.edu/. Applications should include a statement describing the candidate's ability to meet stated qualifications for the position, *curriculum vitae*, and the names and contact information for at least three references. Submit applications (electronic applications preferred) to the chair of the search committee: Michael Ettredge at: mettredge@ku.edu; Professor of Accounting, School of Business, University of Kansas, 1300 Sunnyside Ave., Lawrence, KS 66045-7585. The review of applications will begin immediately and will continue until the position is filled. KU is an Equal Opportunity/Affirmative Action Employer. The University encourages applications from underrepresented group members.

531 WESTERN ILLINOIS UNIVERSITY, Chair and one or two tenure-track faculty positions, Department of Accountancy. Primary need for tenure-track positions is auditing and systems; secondary interest in governmental or fraud is preferred. Other areas may be accommodated. Applicants must have a strong commitment to quality teaching, research, and service, as well as excellent communication skills. The auditing/systems position potentially carries The Sam Oliva Research Fellowship. The College of Business and the Department of Accountancy are accredited by AACSB International. A competitive salary and an excellent benefits package are offered. Screening begins immediately and will continue until the position is filled. Beginning dates are negotiable and subject to funding. For complete details, please see: http://www.wiu.edu/employment or contact JA-Elfrink@wiu.edu. Western Illinois University is an AA/EO employer, and has a non-discrimination policy that includes sex, race, color, sexual orientation, gender identity and gender expression, religion, age, marital status, national origin, disability, and veteran status.

532 SHANGHAI JIAO TONG UNIVERSITY, School of Management invites applications for junior and senior positions in Accounting and Finance. The School plans to grow aggressively in the next several years, and therefore starting dates are flexible. Ranks are open, and salaries are competitive (340K to 600K RMB). Research funds, fringe benefits, and excellent housing will be provided. Applicants should have a Ph.D. in the relevant field by the time appointment starts. Applications for a junior position should demonstrate promise for great scholarship and teaching (undergraduate, M.B.A./E.M.B.A., and Ph.D.); and for senior positions, excellent records of research and teaching are required. The School is located in a prime area in Shanghai that provides excellent living conditions. Please send a cover letter, *curriculum vita*, samples of publications, and three letters of recommendation to: Professor Tian Peng at: ptian@sjtu.edu.cn.

533 RIDER UNIVERSITY, Accounting Department, College of Business Administration invites applications for a tenure-track position beginning in Fall 2010 at the rank of Assistant Professor. For more information on this position and for application instructions, please visit our website at: http://www.rider.edu/hr. Position #: 101115. AA/EOE.

534 BAYLOR UNIVERSITY, Department of Accounting and Business Law invites applications for one anticipated tenure-track position for Fall 2010. Applicants at all academic ranks with teaching and research areas in financial accounting, managerial accounting, tax, or auditing will be considered. A Ph.D. in accounting is required, although ABD candidates will be considered at the Assistant level. Candidates should demonstrate excellent teaching potential and an ability to establish and maintain an ongoing program of substantive research. Professional certification is preferred. Applicants must demonstrate research and teaching excellence. Applications, including *curriculum vitae*, three letters of reference, working paper or defended dissertation proposal, and evidence of an active Christian faith commitment should be sent to: Dr. Charles E. Davis, Department of Accounting and Business Law, Baylor University, One Bear Place, #98002, Waco, TX 76798-8002; Email: Charles_Davis@baylor.edu. Applications will be reviewed beginning October 15, 2009 and will be accepted until the position is filled. Baylor University is a Baptist university affiliated with the Baptist General Convention of Texas. As an Affirmative Action/ Equal Employment Opportunity Employer, Baylor University encourages minorities, women, veterans, and persons with disabilities to apply.

535 SHIPPENSBURG UNIVERSITY invites applications for tenure-track Accounting faculty position, effective January or August 2010. Rank commensurate with qualifications including prior teaching experience and published accounting research. Earned Ph.D. or D.B.A. in Accounting from AACSB institution is preferred. ABDs and advanced degrees in related discipline will be considered if applicant has a CPA or CMA certification with demonstrated research interest in Accounting. Send letter of application, current *curriculum vitae*, undergraduate and graduate transcripts, and contact information for three references to: Dr. Scott N. Cairns, College of Business, Shippensburg University, 1871 Old Main Drive, Shippensburg, PA 17257-2299. Applications are reviewed upon receipt as part of an ongoing search and will continue to be reviewed until the position is filled. Shippensburg University is an Equal Opportunity Employer. Individuals from traditionally underrepresented populations are encouraged to apply. Visit http://www.ship.edu/HR/index.html for additional requirements and qualifications.

536 CALIFORNIA STATE UNIVERSITY, DOMINGUEZ HILLS, invites applications for one full-time tenure-track position at the rank of Assistant, Associate, or Professor in Accounting in the Department of Accounting and Finance, College of Business Administration and Public Policy. The appointment will be effective spring semester 2010 or fall semester 2010. Appointment and start date are subject to the availability of funds. Required: Ph.D./D.B.A. in business administration with a major or concentration in accounting. ABDs and candidates with terminal degrees from related fields are encouraged to apply. A record of or potential for effective teaching and research and publication in the field of accounting. Preferred: Expertise and/or the promise thereof to teach M.B.A. and undergraduate online accounting courses. Relevant practitioner experience and the ability to interact effectively with accounting professionals. Ability to build academic programs. To ensure full consideration, applications should be submitted no later than September 15, 2009. Thereafter, review of applications will begin and will continue until the position is filled. All applications must be submitted through the online application process as follows and PDF document format for required attachments is preferred. Go to the Employment Opportunities web page at CSUDH (http://www.csudh.edu/employment/). Under "New and Returning Applicants" select view Job Opportunities. Login if you have established a username and password already. If not, register and create a username and password. Select to apply with resume. Upload one file containing *c.v.*, letter of interest, three letters of reference, and selected teaching evaluations. Provide contact information and answer questions. Click "Submit." Address letter of interest to: Dr. Mohamed El-Badawi, Chair of the Department of Accounting and Finance, College of Business Administration and Public Policy, California State University, Dominguez Hills, 1000 E. Victoria St., Carson CA 90747; Phone: (310) 243-3571; Email: melbadawi@csudh.edu. Applicants may also contact Dean James T. Strong regarding the position (Email: jstrong@csudh.edu; Office: (310) 243-3548; Cell: (310) 505-2784). Please contact Chair El-Badawi or Dean Strong for an informational interview.

537 CONCORDIA UNIVERSITY invites applications for one tenure-track position in Accountancy. Subject to budgetary approval, we anticipate filling this position, normally at the rank of Assistant Professor, for July 1, 2010. Appointments at a more senior level may also be considered. Review of applications will begin immediately and will continue until the position is filled. To be qualified, applicants must have a doctoral degree (completed or near completion) with concentration in Accounting, and a relevant research and teaching record. The Department of Accountancy is one of five departments at the John Molson School of Business (AACSB-accredited). The department has nearly 30 full-time faculty members, is one of the largest in the country, and has an established record of excellence in professional accounting education and academic research. Please send your application package as soon as possible (i.e., cover letter, resume, statement of research interests and teaching philosophy, teaching evaluations, official Ph.D. transcript, working papers, and three sealed reference letters), but no later than February 1, 2010, to: Dr. Dominic Peltier-Rivest, Chairman, Department of Accountancy, John Molson School of Business, Concordia University, 1455 de Maisonneuve Blvd. West, Office MB 14-215, Montreal, Quebec, Canada, H3G 1M8; Email: drivest@jmsb.concordia.ca. All qualified candidates are encouraged to apply. However, in accordance with Canadian immigration requirements, priority will be given to Canadian citizens and permanent residents of Canada. Concordia University is committed to employment equity.

538 CONCORDIA UNIVERSITY invites applications for a senior Canada Research Chair in Accountancy. To compete for a Canada Research Chair award, the successful candidate is expected to have an outstanding and innovative research program, and the ability to attract research grants and excellent Ph.D. students. Applicants must have a Ph.D. in Accountancy or a related discipline, a superior record of publications, a proven track record of successful research program leadership and broad-based international exposure. Candidates for Tier I Canada Research Chair must be Full Professors or Associate Professors who are expected to be promoted to the Full Professor level within one or two years of the application. Tier I chairs are tenable for seven years and are renewable. Applications must include a *cv*, a reasoned letter in which you discuss your qualifications for this Chair including your three most significant scholarly achievements, a full and detailed seven-year research plan, and the names, addresses, and current contact information of three referees. Applications will be reviewed as they arrive and the search will continue until the position is filled. All applications should reach the department no later than February 1st, 2010. The candidate nominated by the University will have his or her application submitted for approval to the CRC Secretariat. This position is subject to budgetary approval and will be filled upon the successful completion of the CRC nomination and approval process. Further details about the CRC Program, as well as the chair selection criteria, can be found on the CRC website at: http://chairs.gc.ca. Your application should be sent to: Dr. Dominic Peltier-Rivest, Chairman, Department of Accountancy, John Molson School of Business, Concordia University, 1455 de Maisonneuve Blvd. West, Office MB 14-215, Montreal, Quebec, Canada, H3G 1M8; Email: drivest@jmsb.concordia.ca. All qualified candidates are encouraged to apply. However, in accordance with Canadian immigration requirements, Canadian citizens and permanent residents of Canada will be given priority. Concordia University is committed to employment equity.

539 NORTHERN KENTUCKY UNIVERSITY, Department of Accountancy's Haile/US Bank College of Business invites applications for a tenure-track position at the Assistant or Associate Professor level beginning in August 2010, with primary teaching responsibility in the area of auditing. Requirements include a Ph.D. in accounting, commitment to teaching excellence, and demonstrated ability to do quality academic research. Professional certification and practical accounting experience preferred. Salary and benefits are AACSB competitive. NKU is located in the beautiful rolling hills of northern Kentucky just five miles from downtown Cincinnati. The College of Business is AACSB-accredited and offers both undergraduate and graduate degrees including a Master's of Accountancy. Additional information is available at: http://www.nku.edu/~cob. Qualified candidates should submit a letter of application, current resume, and contact information for three references electronically to: Dr. Carol Lawrence, Chair, Department of Accounting, at: lawrencec5@nku.edu. Northern Kentucky University is an Affirmative Action/Equal Opportunity Employer.

540 INDIANA UNIVERSITY-PURDUE UNIVERSITY FORT WAYNE (http://www.ipfw.edu) seeks a tenure-track Assistant Professor for August 2010 whose primary teaching interest is in Managerial Accounting. A doctorate in accounting is required. The successful candidate must demonstrate a commitment to high-quality teaching, intellectual contribution, and service. IPFW has recently been reaffirmed as an AACSB-accredited institution. Applicants should send a cover letter, *curriculum vita*, and a list of three references to: Dr. Kathy Pollock, Chairperson, Indiana University-Purdue University Fort Wayne, Department of Accounting and Finance, 2101 E. Coliseum Blvd., Fort Wayne, IN 46805; or Email: pollockk@ipfw.edu. IPFW is an Equal Opportunity/Equal Access/Affirmative Action University.

541 UNIVERSITY OF PENNSYLVANIA, Wharton School's Accounting Department is seeking applicants for a full-time, tenure-track faculty position at any level: Assistant, Associate, or Full Professor, in Financial, Managerial, and Tax Accounting. Demonstrated ability or potential for high-quality research and teaching are important. Applicants must have a Ph.D. (expected completion by June 30, 2011 is acceptable) from an accredited institution and have an outstanding research record or potential. The appointment is expected to begin July 1, 2010. Applications, including cover letter, *curriculum vitae*, and recent publications or working papers, must be submitted via our secure website: http://accounting.wharton.upenn.edu/. Junior-level candidates should also include contact information for up to three references who will be asked to submit letters of recommendation. The deadline for applications is January 15, 2010. Contact: Professor Brian Bushee, Accounting Department, The Wharton School, University of Pennsylvania, 1300 Steinberg-Dietrich Hall, Philadelphia, PA 19104-6365. The University of Pennsylvania values diversity and seeks talented students, faculty, and staff from diverse backgrounds. The University of Pennsylvania is an Equal Opportunity/Affirmative Action Employer. Women, minority candidates, veterans, and individuals with disabilities are strongly encouraged to apply.

542 DREXEL UNIVERSITY, LeBow College of Business, Department of Accounting and Taxation invites applications for two positions (one Assistant Professor and one open-rank position) starting September 1, 2010. Assistant Professor applicants should have received their doctorates before the beginning of the contract period. Candidates for senior positions should demonstrate a strong record of research and evidence of teaching excellence. The preference for the Assistant Professor position is in Accounting Information Systems or Managerial Accounting. The preference for the open-rank position will be Auditing or Managerial Accounting. The LeBow College of Business offers an undergraduate accounting degree, an M.S. in Accounting, and a Ph.D. with an accounting emphasis. The Department has a rich tradition of interfacing with the Philadelphia business community through our well-established undergraduate co-op program and an active Accounting Advisory Board. Positions will remain open until filled. Send a letter of application, *vita*, and other supporting materials electronically to: http://www.drexeljobs.com. Drexel University is an Equal Opportunity/Affirmative Action Employer and encourages applications from women, members of minority groups, disabled individuals, and veterans.

543 THE PENNSYLVANIA STATE UNIVERSITY, School of Graduate Professional Studies at Great Valley invites applications for a tenure-track position in accounting for its AACSB-accredited evening and weekend M.B.A. and weekend Master of Finance programs. Located 25 minutes northwest of Philadelphia, the Penn State School of Graduate Professional Studies at Great Valley (http://www.sgps.psu.edu) is the nation's first university facility permanently housed in a corporate park, where it stands among world-class corporate neighbors. It was created solely to accommodate the busy schedules of full-time professionals and emphasizes the link between academic theory and professional practice. Excellence in both teaching and research is expected and supported through technology and other services. Faculty members also advise students and fulfill program, campus, and University service assignments. Candidates must hold an earned doctorate in accounting. The position will be available for the Fall 2010 semester. Qualified applicants should send a letter of application, *curriculum vitae*, and the names, and contact information of three references to: Ms. Elizabeth Keane-Carney, Attn: Accounting Faculty Position, Penn State Great Valley, 30 East Swedesford Road, Main Bldg., Malvern, PA 19355; Email: ekc1@psu.edu. Penn State is committed to Affirmative Action/Equal Opportunity and the diversity of its workforce.

544 LINGNAN UNIVERSITY is inviting applications for the following positions (tenable from mid-August 2010 or by mutual agreement thereafter): (1) Professor/Associate Professor/Assistant Professor Department of Accountancy (Post Ref.: 09/100/AAA). The appointee should be able to teach in at least two of the areas: Financial Accounting, Auditing, Accounting Information Systems, and Managerial Accounting. CPA qualification and demonstrated research strength are advantages. (2) Professor Department of Computing and Decision Sciences (Post Ref.: 09/101/AAA). Applicants should have teaching and research experience in two or more of the areas: Information Systems, Operations Management, Management Science and Statistics, Accounting Information Systems, and Forensic Accounting. Applicants for Professor should have a current and substantial research record and teaching experience with significant scholarly publications in leading journals. Applicants for Associate Professor should have a Ph.D. degree and at least 5 years of relevant experience and a sound record in research and scholarly activities. For appointment as an Assistant Professor, applicants should have a Ph.D. degree in the relevant discipline with teaching experience and a good research record. Professor: around HK$75,500 to HK$101,400; Associate Professor: around HK$61,600 to HK$81,200; Assistant Professor: around HK$45,300 to HK$71,400. The rank and remuneration will be commensurate with qualifications and experience. Fringe benefits are included. Please forward your dossier (including reference letters) and a form (R1), obtainable at: http://www.LN.edu.hk/hr/appform.htm, to: Human Resources Office, Lingnan University, Tuen Mun, Hong Kong; or Fax: (852) 2891-5782; or Email: recruit@LN.edu.hk. Please specify the post you are applying for and quote the Ref number of the post.

545 KUTZTOWN UNIVERSITY, College of Business has a stated mission of excellence in teaching. The Department of Accounting and Finance is announcing two tenure-track faculty positions to teach Accounting, beginning in Fall 2010. Required teaching load is 12 semester hours each semester. Courses will include a combination of lower-level and upper-level Accounting courses, and possibly M.B.A. The Kutztown University faculty are represented by a collective bargaining agreement that regulates the salaries of all faculty. Doctorate in Accounting or related area is preferred. ABD in Accounting with expected completion within one year will be considered. Interested applicants should respond to the address below and include the following information: (1) An up-to-date *curriculum vitae*; (2) Transcripts for all academic work completed (official transcripts are required before hiring); (3) Names and contact information of at least three references (letters from the referees will be required before interview); and (4) A cover letter which clearly outlines the applicant's teaching interests and research interests. This information should be sent to: Norman Sigmond, Chairperson, Accounting Faculty Search, Accounting and Finance Department, Kutztown University, Room 131, DeFrancesco Hall, Kutztown, PA 19530. Review of applications will begin on September 1, 2009 and will continue until the positions are filled.

546 SAN JOSE STATE UNIVERSITY invites applications for tenure-track faculty positions starting in August 2010. The available accounting positions are at the Associate or Assistant Professor rank for Audit and at Full, Associate, or Assistant rank, or a temporary Visiting Professor for Taxation. The candidate must have demonstrated versatility in subject matter competencies and pedagogy. The teaching assignment is at both graduate and undergraduate levels including traditional and accelerated formats. The positions are subject to budget approval. The minimum credential requirement is ABD for tenure-track and Visiting appointments. Strong preference will be given to candidates who hold a Ph.D. from an AACSB-accredited university. The application deadline is May 31, 2010 or until the position is filled. Send application letter, *vita*, teaching evaluations, and three letters of reference to: Dr. Janis Zaima, Chair, Accounting and Finance Department, College of Business, San Jose State University, One Washington Square, BT850, San Jose, CA 95192-0066; Phone (408) 924-3490; Email: zaima_j@cob.sjsu.edu. SJSU is an Equal Opportunity/Affirmative Action Employer committed to the core values of inclusion, civility, and respect for each individual.

547 SAN JOSE STATE UNIVERSITY, position for Summer 2010. Specialization: Graduate Level Taxation, (subject to budget approval). Qualifications: The applicant must have experience teaching graduate level tax courses, high-quality teaching skills, and meet SJSU's definition of academically qualified (AQ) per AACSB-accreditation standards (see http://www.sjsu.edu/facultyaffairs/jobs/ for details; qualifications include possessing a Ph.D., J.D. or M.S.T., and at least two peer-reviewed journal articles in appropriate publications since 2006). Experience teaching S corporations, estate and gifts, IRS practice and procedure, and/or taxation of tax-exempt organizations. Responsibilities: Prepare and teach one or two M.S.T. courses in summer 2010 (mid-June to mid-August 2010). Class meetings are 4 hours and are scheduled for one evening per week for eight weeks + two Saturday mornings. Classes meet in San Jose. Must adhere to scheduling, grading, and other instructional policies of SJSU and the College of Business. Course syllabus needed by March 2010. Salary Range: Commensurate with qualifications and experience. A research grant may also be available in addition to teaching (depending on timing, interest, research area, and co-authoring possibilities). Salary information is available at http://www.fa.sjsu.edu/. Employment Period: Class session runs approximately mid-June 2010 to mid-August 2010 Eligibility: Employment is contingent upon proof of eligibility to work in the United States. Application Procedures: For full consideration send a letter of application, *curriculum vitae*, statement of relevant teaching experience, and at least two letters of reference with contact information to: Annette Nellen, Director, MST Program, San Jose State University, One Washington Square, San Jose, CA 95192-0066; or Email: anellen@sjsu.edu. Questions: Contact Professor Annette Nellen at anellen@sjsu.edu or (408) 924-3508.

548 NATIONAL CHENGCHI UNIVERSITY, Department and Graduate School of Accounting, College of Commerce, Taipei, Taiwan, http://acct.nccu.edu.tw. The College of Commerce of National Chengchi University is fully accredited by AACSB and aims to further enhance its quality of teaching and research. The Department of Accounting invites applications for faculty positions at all levels. Candidates majoring in Financial Accounting, Managerial Accounting, Auditing, Taxation, and Accounting Information Systems are welcomed. Applicants must have a Ph.D. degree and strong research records. Newly appointed faculty members would be awarded a generous grant, a salary premium offered to courses taught in English, lower teaching load, and research and teaching awards. Candidates with outstanding research records could be offered a chair Professorship at the College level. Interested candidates should send *curriculum vitae* and three recommendation letters to: Professor and Chair Jia-Lang Seng at: seng@nccu.edu.tw; and Ms. Chunfen Liu at: chunfen@nccu.edu.tw; Phone: (886) 2 2939 3091, ext. 87042.

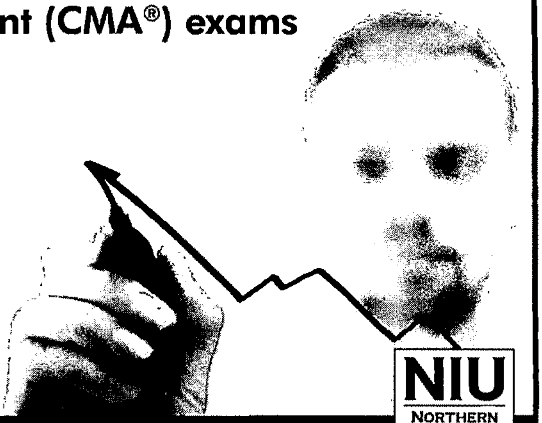
549 TEXAS STATE UNIVERSITY SAN MARCOS invites applications for the position of Chair of the Department of Accounting in the McCoy College of Business Administration. Candidates must have a doctorate in accounting, business administration or a J.D. with an L.L.M., administrative skills needed to lead a department in a large public university, and a record of excellence in scholarship, teaching, and service commensurate with appointment at the Professor rank. For more information, see http://accounting.mccoy.txstate.edu. To apply, please submit a completed Faculty Employment Application available at: http://www.txstate.edu/academicaffairs/faculty%20employment%20application.doc, *curriculum vitae*, college transcripts, three references (including name, address, telephone, email address), and an application letter to: Dr. Ray Fisk, Attention: Search Committee, Department of Marketing, McCoy College of Business, Texas State University San Marcos, 601 University Drive, San Marcos, TX 78666. Electronic materials should be submitted to: rf17@txstate.edu.

550 TEXAS STATE UNIVERSITY SAN MARCOS invites applications for the position of Assistant/ Associate Professor, tenure- or clinical-track, Department of Accounting in the McCoy College of Business Administration. Candidates must have a doctorate in accounting, business administration or a J.D. with an L.L.M., or compatible degree, and potential or record of excellence in scholarship, teaching, and service. For more information, see http://accounting.mccoy.txstate.edu. To apply, please submit a completed Faculty Employment Application available at: http://www.txstate.edu/academicaffairs/faculty%20employment%20application.doc, *curriculum vitae*, college transcripts, three references (including name, address, telephone, email address), and an application letter to: Dr. Lu Montondon, Search Committee, Department of Accounting, McCoy College of Business, Texas State University San Marcos, 601 University Drive, San Marcos, TX 78666. Electronic materials should be submitted to: lm03@txstate.edu.

551 UNIVERSITY OF ILLINOIS AT URBANA–CHAMPAIGN, Department of Accountancy is recruiting faculty for full-time appointments at all ranks. Candidates at advanced ranks must possess a doctoral degree, and candidates for Assistant Professor positions must possess or be near completion of a doctoral degree. Applicants for senior positions should have a strong research record and the ability to participate in doctoral education and curriculum development at the undergraduate and graduate levels. In addition, non-tenure-track Lecturer and Visiting positions are available. Applicants should submit a resume that lists research publications and courses taught. Compensation is competitive and will be commensurate with the appointment level and applicant's record. Please create your candidate profile at http://jobs.illinois.edu and upload the required documents, including cover letter, *curriculum vitae*, teaching statement, and the names and email addresses or phone numbers of three professional references. Proposed starting date is June 16, 2010 or August 16, 2010. For further information please refer to: http://www.business.illinois.edu/accountancy/recruit.html; or email: accyrecruiting@business.illinois.edu. Contact: Recruiting Chair, Department of Accountancy, 360 Wohlers Hall, College of Business, University of Illinois, 1206 S. Sixth Street, Champaign, IL 61820. The University of Illinois is an Affirmative Action/Equal Opportunity Employer. The administration, faculty, and staff embrace diversity and are committed to attracting qualified candidates who also embrace and value diversity and inclusivity.

552 UNIVERSITY OF MACAU, Faculty of Business Administration (FBA) invites applications for the Associate/Assistant Professor academic position in the area of Accounting. The University of Macau is a government-funded institution located in the Macao Special Administrative Region of the People's Republic of China. Qualifications: Appointee is required to teach courses at the undergraduate and graduate levels and to provide M.B.A. dissertation supervision. The candidate is expected to teach and do excellent research in the following area(s): Financial Accounting, International Accounting, Auditing, and Management Accounting. Position and Taxable Annual Salary: The position offered and taxable salary level shall be determined according to the appointee's academic qualifications, current position, and professional experience. For details about the "Terms of Academic Appointment," please refer to: https://isw.umac.mo/recruitment/showTermsAppoint.do. Associate Professor: MOP660,800–MOP743,400; Assistant Professor: MOP536,900–MOP619,500, (U.S.$1 = approx. MOP8). Selected candidate is requested to assume duty in September 2010.

553 MURRAY STATE UNIVERSITY, Department of Accounting invites applications for a tenure-track Associate Professor or Professor position as an endowed professorship. Anticipated opening, beginning August 15, 2010. Candidates must hold a doctoral degree from an AACSB-International accredited program with a residence requirement. Professional certification and substantive professional experience are required. Candidates also should have a research record commensurate with employment at the Associate Professor or Full Professor level. The position will pay an AACSB-competitive salary plus a substantial stipend. All teaching areas will be considered. The successful candidate will teach courses at the undergraduate and graduate level and publish in peer-reviewed accounting journals. Responsibilities will include developing new and innovative instructional and learning experiences for students. The complete job announcement and requirements can be found at: http://www.murraystate.edu/indir/hr/jobrpt.htm or the AAA Career Center website. Interested candidates should send a letter of application, *curriculum vita*, graduate transcripts, evidence of teaching and research effectiveness, and three current letters of recommendation to: Chair of Accounting Search Committee, Murray State University, Department of Accounting, 351 Business Building, Murray, KY 42071. Application deadline is October 2, 2009. Women and minorities are encouraged to apply. Murray State University is an Equal Education/Employment Opportunity, M/F/D, AA employer.



There's Always Room for Professional Growth

Prepare for the **Certified Internal Auditor (CIA®)** or the **Certified Management Accountant (CMA®) exams** starting in August.

For more information, visit our web site at www.cia.niu.edu or www.cma.niu.edu

NIU Professional Development

NIU
NORTHERN ILLINOIS UNIVERSITY

For focused results in accounting and tax research, start here.

While the world of accounting information continuously transitions, professionals, educators, and a growing number of accounting students look to reliable resources for researching accounting trends, rules, issues, and history pertaining to the future of accounting...

ProQuest Accounting & Tax, our premier accounting research database, makes research time more effective with centralized access to top accounting journals, reference reports, proceedings, dissertations, and more.

Researchers can instantly cross-search over 2,300 publications, with hundreds in ongoing full text, plus dissertations and working papers covering accounting and tax topics.

ABI/INFORM® subscribers can enhance their business database library with ProQuest Accounting & Tax, which provides a vast array of additional full-text accounting journals, trade publications, dissertations, and working papers not included in ABI/INFORM.

ProQuest Accounting & Tax includes:

- Highly ranked global and scholarly journals, including all journals from the American Accounting Association
- Accounting, tax management, and auditing trade publications from top publishers such as CCH, Aspen, Thomson Professional and Regulatory Services, Source Media, Tax Management Inc., and Euromoney
- AICPA publications
- International Federation of Accountants (IFAC) standards
- Tax law reviews from top institutions such as New York University, Stanford, Vanderbilt, and University of Virginia
- Tax law and financial law publications from noted sources such as the American Bar Association and American Law Institute
- Global accounting and tax journals from the UK, Canada, Australia, France, Spain, Ireland, and more
- Financial advising, governance, and risk management publications from Incisive Media, CFA Institute, and Association for Financial Counseling and Planning
- Accounting history from publications such as *The Accounting Historian's Journal*
- Over 1,500 accounting dissertations and more than 7,800 working papers from SSRN, plus select conference proceedings

ProQuest Accounting & Tax

Visit **www.proquest.com/go/accounting**
today to learn more, and sign up for a free trial!

ProQuest®
Start here.

THE ACCOUNTING REVIEW
ISSN 0001-4826

PERIODICALS POSTAGE PAID
SARASOTA FLORIDA
AND ADDITIONAL MAILING OFFICE

POSTMASTER
Send Address Changes To:

AMERICAN ACCOUNTING ASSOCIATION
5717 BESSIE DRIVE
SARASOTA, FL 34233

34314 REVIEW 1
CAL UNIV
1 KYD STREET
SHOP NO 5 PALACE COURT
INTERNATIONAL BOOK HOUSE PVT LTD
KOLKATA 700 016
INDIA

3071